Vol. 53, No. 4 July, 1946

Psychological Review

EDITED BY
HERBERT S. LANGFELD
PRINCETON UNIVERSITY

CONTENTS

Clarence Stone Yoakum: 1879–1945: W. B. PILLSBURY 195

Paired Comparisons and the Logic of Measurement: HAROLD GULLIKSEN 199

Repetition and Learning. III. Memory and Thought: THOMAS W. COOK 214

The History of Dream Theory: HAROLD GRIER MCCURDY 225

Age as a Factor in the Recall of Interrupted Tasks: R. NEVITT SANFORD 234

The Meaning of Explanation: DAVID L. MILLER 241

Operationism and Psychological Theory: W. C. H. PRENTICE 247

PUBLISHED BI-MONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1227 NINETEENTH STREET, N. W., WASHINGTON 6, D. C.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

PUBLICATIONS OF

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.
Subscription: $5.50 (Foreign, $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

JOHN E. ANDERSON, *Editor*
University of Minnesota

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears bi-monthly beginning in January. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

Subscription: $7.00 (Foreign, $7.25). Single copies, $1.25.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor, on Leave,* FRANCIS W. IRWIN, *Acting Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears bi-monthly (beginning 1944), one volume per year, each volume of six numbers containing about 520 pages.

Subscription: $7.00 (Foreign, $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 350 pages.
Subscription: $6.00 per volume (Foreign, $6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 400 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.
Subscription: $5.00 (Foreign, $5.25). Single copies, $1.50.

JOURNAL OF APPLIED PSYCHOLOGY

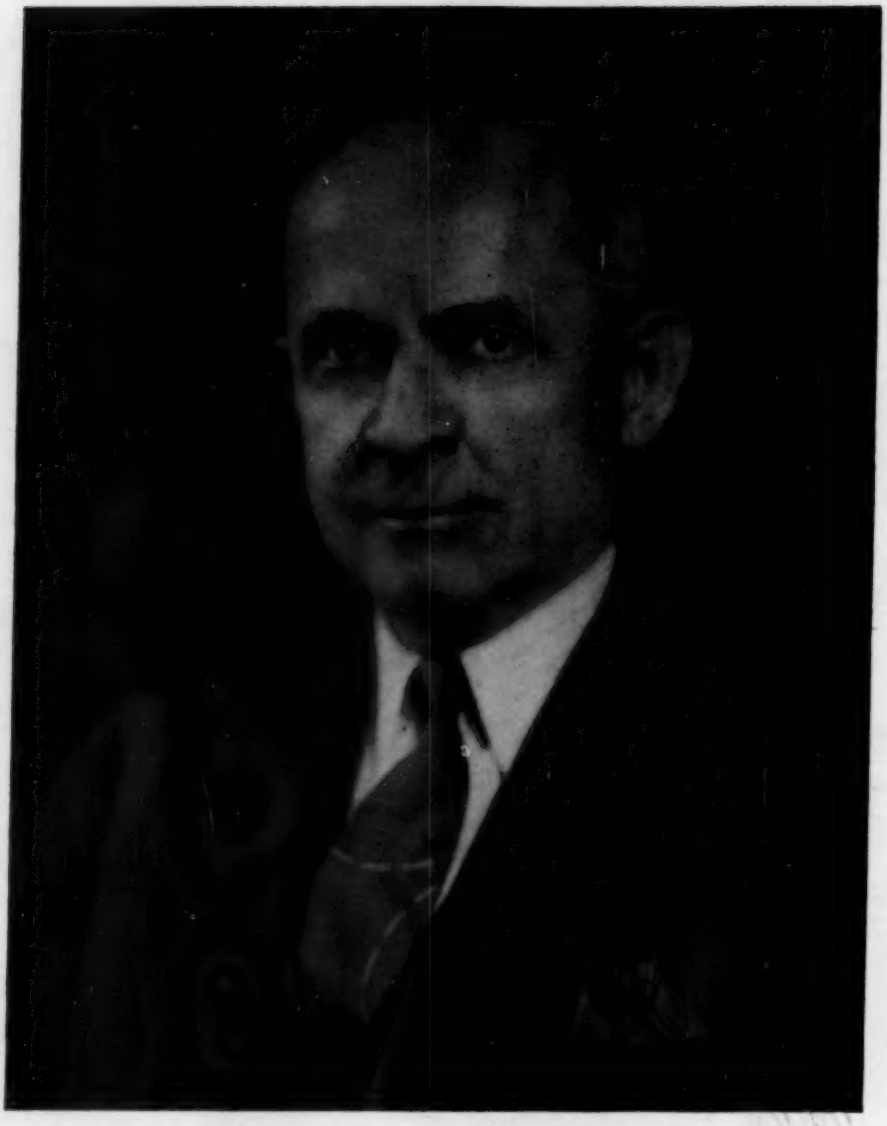
DONALD G. PATERSON, *Editor*
University of Minnesota

Covers the applications of psychology in business, industry, education, etc. Appears bi-monthly, February, April, June, August, October, and December.
Subscription: $6.00. Single copies, $1.25.

Subscriptions, orders, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1227 NINETEENTH STREET, N. W., WASHINGTON 6, D. C.



VOL. 53, NO. 4 JULY, 1946

THE PSYCHOLOGICAL REVIEW

CLARENCE STONE YOAKUM

1879–1945

Yoakum was a psychologist who worked in many fields of the science and also in administration. Three relatively definite phases of his career can be distinguished. As a graduate student at Chicago and professor at the University of Texas he investigated human fatigue and animal learning, with an excursion or two into the borders of the abnormal. The First World War turned him to the field of testing. After the war he transferred his efforts in testing to personnel and related fields of industrial psychology. After a few years, these shifted again to problems of education and administration. His interests and activities showed a wider range than those of most men.

Clarence Stone Yoakum was born in Lawrence County, Kansas, on January 11, 1879. He was an A.B. of Campbell College, Kansas, in 1901, and continued there for two years after graduation as instructor in psychology and education. The following years he was an instructor in mathematics and science in Hiawatha Academy, Kansas. He was appointed to a fellowship in the University of Chicago which he held for two years, receiving the doctorate in 1908. Immediately following he was appointed to a professorship and chairmanship of the Department of Psychology at the University of Texas. Here he remained until the outbreak of the war in 1917, when he entered the psychological service of the Army. During the war he was advanced in rank from First Lieutenant to Major, and in function from testing at Camp Lee to supervisor of testing work at all Army camps.

In 1919 he went to the Carnegie Institute of Technology at Pittsburgh where he joined a group of men who had also been working with tests in the Army program. Under Bingham they adapted the knowledge and skills they had acquired to related problems in industry. Yoakum held the post of Professor of Applied Psychology and Director of Personnel Research from 1919 to 1924. In 1919 he married Louise Branch Storey of Austin, Texas, and they entered upon the life at Pittsburgh together. At this time he became interested in the application of the results of studies of the individual's interest to a determination of the most suitable profession. Although Yoakum himself published relatively little, Freyd and others give credit to his direction in the obtaining of their results. While it lasted, this group was a vital influence in industrial psychology.

In 1924 Yoakum was appointed Professor of Personnel Management and Director of Personnel Research in the School of Business Administration at the University of Michigan. In 1927 he added to his duties the direction of a bureau of educational research for the university as a whole. In 1929 he was

called to Northwestern University as Dean of the College of Liberal Arts. After one year he was called back to the University of Michigan as Vice-President, with his main duties the direction of educational research and as consultant in matters of university personnel. In 1935 he was made Dean of the Graduate School in addition to his other work. Shortly after he assumed these duties, the Rackham Foundation made available funds for a building and for a research fund whose direction fell largely to his charge. This fund was used in part for research in human relations, tests and a speech clinic, among other activities.

Yoakum's publications were also varied in scope, corresponding to his teaching and administrative activities. His first work was his thesis 'An experimental study of fatigue,' published as a *Psychological Review Monograph* in 1909. The method used was unique. The students were first drilled to make series of groups of taps or lip movements in a definite pattern until the process became automatic. The movements were recorded on a kymograph. Then they were asked to add a given number of units to each group in the pattern or to multiply each group by a number, and this changed pattern was repeated. It was assumed that a change in rate of movement or the number of errors would give an indication of fatigue. This expectation was not confirmed. Both rate and errors varied little and irregularly. The only marked indication of fatigue was an accumulation of errors in groups here and there in the work, and, in two individuals, a complete collapse which prevented further recording. The periods of accumulated errors distributed themselves irregularly over the course of the work. They did not seem to be related to the amount of work done. Yoakum suggested that they might be related to physiological rhythms, but did not conjecture what rhythms were involved, and made no attempt to correlate them in the records with known rhythms.

Introspections were made that indicated some relation between the imagery used in controlling the movements and the development of errors. Failure of imagery or periods of distraction accompanied some of the groups of errors. Main importance in theory was placed upon the movements. Fatigue sensations were mostly ascribed to muscles, and usually to muscles that were not supposed to be used in the movements of tapping but to remote and more general groups. Some statements seem to anticipate the behavioristic theory that was to be formulated a few years later. Yoakum made some remarks later that indicated his acceptance of that view but never published a definite statement of his beliefs on the subject.

A study of the learning process in squirrels also dates from the Chicago period. This is a brief paper whose main contribution is that squirrels learn by a trial-and-error process. They differed from white rats mainly in never permitting themselves to be handled, in having their learning processes inhibited readily by emotional reactions and being subject to confusion of the method being learned with other similar experiments. The same paper gave results on temperature discrimination. The method was rather rough and shows only that differences of 25° may be discriminated.

While Professor at the University of Texas Yoakum published a report on the complication experiment and two studies of imagery and memory. The brief statement of the complication experiment suggested that the after-image of the pointer might disturb the results. The results themselves are not different

from those reported by other workers.

The more interesting of the memory studies described a case of exceptional memory marked by ability to recall capital letters presented twelve at a time, three rows of four each, even after the lapse of considerable periods. The individual was a girl exceptionally competent as a student, but rather unadjusted socially, and tending to the abnormal. It proved that the observer spontaneously connected the letters to be learned in groups of two or more with an incident and would build the incidents into a story. The stories were sometimes drawn from her experience, at other times they were constructions that had become fixed enough to approach the complexes of the abnormal. Analysis of the method explained the unusual accuracy in recall and aided in understanding some of the peculiarities of the student.

Best known of the works with which Yoakum was associated is 'Army mental tests,' in collaboration with Yerkes. After the various committees brought together reports on different phases of the work, Yerkes and Yoakum worked over the results to give an over-all picture of the methods, of the conclusions that might be drawn from them for the use of the army classification and the broader implications for many political, social and educational problems. The early portion of the book gives an outline of the relationship between standing in the tests and various evidence of military value. These showed that rank was closely similar to the standing in the scores. Standing in officers' training schools correlated closely with grades made in the tests. The branches of the service differed greatly in intelligence as measured by the test, although some of the results were unexpected. Medical officers were near the bottom, engineers at the top. The occupations in civil life showed a close relation to the grades made. The material of the book has become so well known that it needs no summary.

For a period after the war Yoakum devoted himself to problems of personnel in industry. His courses at the Carnegie Institute of Technology and his practical work with business concerns put an emphasis upon this field in his writing as well. He studied the methods that might be used in selecting a profession for an individual and the selection and training of employees by business firms. Among several publications on these topics we may mention 'The selection and training of salesmen' in which he was associated with Kenagy. It is a thorough treatment of the whole problem of sales management. Three chapters are devoted to psychological tests and the limits of their practical application. The other chapters cover the organization and management of services and the details of training and supervision. It is definitely a practical book but does contain many references to psychological principles and important statements on psychological problems incidental to the more practical purposes.

Yoakum added to his professorship the duty of directing a bureau of university research in 1927. He had ample funds at his disposal and competent assistants. The duties and facilities gave a direction to the researches of his later life. In the early period he suggested a series of studies on correlations of various criteria used (or indicated) to measure the capacities of students in the university. These studies were continued and extended over a number of years. The results were brought together in 1940 in a study by Yoakum, P. S. Dwyer and Charlotte Horner entitled 'A statistical summary of the records of students entering the University of Michigan as freshmen in the decade 1927–1936.' Hundreds of different cri-

teria were used in the correlations. Most emphasis was given to sixty-four of the measures frequently used as the basis of college entrance. Correlations with university grades range from − 12 for age to 0.46 for some of the tests. The total result indicates most strongly that there is a large element of chance in the selection of individuals for college entrance. This result might be expected, since grades for different years in college correlate for the most part only in the fifties.

The study covered some ten thousand individuals over a decade. It is one of the most comprehensive investigations on record in the number of individuals considered and in the number of measurements and criteria applied. It should be definitive in the subject until new tests are devised. Other studies are in the preliminary stages.

As an administrator Yoakum won the respect of his superiors. In the army he impressed the military with his ability to get things done. His colleagues relied on him to straighten out the work of testing in camps where friction had developed. The positions he held in universities indicate the esteem in which he was held by the men in authority.

W. B. PILLSBURY
University of Michigan

PAIRED COMPARISONS AND THE LOGIC OF MEASUREMENT

BY HAROLD GULLIKSEN

Princeton University

I. Introduction

During the past decade several critiques and surveys of psychological measurement have appeared in the literature (1, 2, 7, 8, 12, 16, 17). Some of these articles have given a detailed discussion of certain methods of psychological measurement and concluded that these methods lack the fundamental characteristics of true measurement (2, 7, 8, 12, 17). These articles, however, have entirely ignored important developments of the last twenty years in the theory dealing with the method of paired comparisons (9, 10, 11, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24). It is important to consider paired comparisons and the law of comparative judgment carefully, when considering methods of psychological measurement, since this method has great flexibility and generality, has a good theoretical foundation, and a complete system of internal checks on itself. The law of comparative judgment, stated by Thurstone (18), provides a basic method, which is free from many of the objections that can be raised against other methods.

The purpose of this paper is to present a somewhat broader definition of measurement than that given by other writers (2, 3, 4, 5, 6, 7, 8), and to show that scales constructed by the method of paired comparisons will satisfy this broader definition of measurement. It will be shown that such scales have a point for point correspondence with certain types of physical measurement, and that such scales have valuable properties not possessed by ordinal or rank order scales.

Since it is generally agreed that psychophysical scales satisfy the requirements of ordinal scales we will first review briefly the requirements of ordinal scales and show in particular that scales constructed by the method of paired comparisons satisfy these requirements. We will then consider the more controversial topic of the relationship of paired comparisons to 'extensive' or 'additive' scales. Since all the writers on the topic of measurement have accepted the requirements for measurement laid down by Campbell (4) we will consider in particular the relation of paired comparison scales to the requirements listed by Campbell.

II. Ordinal Scales

The establishment of an ordinal or rank order scale for a series of objects or 'systems' depends upon the existence of two relationships, of which one is transitive and asymmetric, and the other is transitive and symmetric, as indicated by Campbell (4, p. 5). Such relations as 'brighter than,' 'louder than,' and 'heavier than' are illustrations of transitive asymmetric relations that obtain between sensations, and can be detected by the method of paired comparisons. If the percentage of subjects who judge "A brighter than B" is greater than 50 per cent, then our 'test' shows that A is brighter than B. This relationship is obviously asymmetric, and numerous psychophysical experiments have shown that the relation is transitive. That is to say, if the percentage of judgments "A brighter than B" is greater than 50 per cent and the percentage of judgments "B brighter than C" is greater than

50 per cent, then we also find that the percentage of judgments 'A brighter than C' is greater than 50 per cent.

Such relations as 'equally bright,' 'equally loud,' and 'equally heavy' are illustrations of transitive symmetric relations that obtain between sensations. Two lights are presented and the subjects are asked "Which is brighter?" If a tabulation of the results shows that 50 per cent say "A is brighter than B" and 50 per cent say "B is brighter than A" then the two lights are 'equally bright.' This relationship is clearly symmetrical, so that if the evidence shows that A equals B, then also B equals A. Furthermore, our general knowledge of results in psychophysics tells us that if another light C is found, such that we get 50 per cent judgments "B brighter than C" and 50 per cent "C brighter than B," then within the limits of experimental error we will find that 50 per cent will judge "A brighter than C" and 50 per cent will judge "C brighter than A." Thus the relation will be experimentally demonstrated to be transitive, since if A equals B and B equals C, then A equals C. It should be noted here that in psychophysics, just as in physical measurement, one must use very precise methods in order to verify the transitive property of the relation 'equals.' It is easily possible to select a series of, let us say, six sensations (A, B, . . . F), such that the judgments would be "A equals B," "B equals C," etc. to "E equals F," and then to find "A brighter than F." The same sort of difficulty is encountered in physical measurement of, let us say, a series of weights, each differing from the other by a fraction of an ounce. In determining equality it is necessary to make use of extremely precise measurements if one is to avoid this difficulty. In the case of sensations, the method of paired comparisons furnishes the best method of obtaining precision, since an indefinitely large number of judgments can be obtained, and two sensations are defined as equal only when exactly fifty per cent of the judgments are "A heavier than B" and the other fifty per cent are "B heavier than A." If this is true for several hundred judgments, then the transitive property of the relation 'equals' will doubtless be verified for the method of paired comparisons. Also, just as one must be careful to rule out the effects of friction in a balance, so one must adopt precautions against time, space, and other sorts of error in psychophysics. With some precautions such as these regarding precision of method, it is generally agreed that sensations and other psychological material exhibit these two types of relationships (transitive asymmetric and transitive symmetric) and hence that ordinal scales can be established.

In summary of this first point then, it can be seen that psychophysics provides a test for the relationships 'equals' and 'greater than' and that the first of these relationships is symmetric and transitive, while the latter is asymmetric and transitive. Our experience with psychophysical scales is sufficient to indicate that when one makes due allowance for experimental error, then these relationships also fulfill the 'Axioms of order.' Given a set of N objects we can rank them in order and assign a number (from 1 to N) to each of the objects.

Such a scaling method, however, is defective in that the number assigned to an object depends entirely upon the particular choice of the N objects to be scaled. A given object can have the number '3' in one series and the number 'N' when it is regarded as a member of another series. New objects not in the original set have only an approximate position in the scale. They may be located as between two members of the scale, or as being outside of the scale at

one end or the other. If two different scientists set out to scale the same attribute, the results obtained by one could not be readily translated into the terms of the other, unless, of course, they had previously reached complete agreement on the exact set of N objects to be scaled. Furthermore, there is no arithmetical operation that can be performed with the assigned numbers (1 to N) that has any parallel in an operation which can be performed with the objects. For example, if we take the object number '2' and object number '5' in a series we cannot, by using these two objects, perform any operation the results of which are indicated by '3' $(5 - 2)$ or by '7' $(5 + 2)$, etc. That is to say, there is no arithmetical operation on ordinal numbers that has a parallel in an experimental operation with the corresponding objects. There is, of course, the trivial exception that the difference between two numbers is one greater than the number of objects on the arbitrary scale lying between the two objects with these numbers. For example, between object '6' and object '2' there are three (or $6 - 2 - 1$) objects. However, counting the number of objects lying between objects '6' and '2' is not *an operation* WITH *the objects '6' and '2.'* The 'extensive' or 'additive' type of scale does not have these limitations. We turn next to a consideration of this type of scale.

III. EXTENSIVE SCALES

Additive magnitudes are divided into three types: fundamental, derived, and defined (4, pp. 14 and 94). The terms defined and derived are applied to magnitudes which are measured only as some function of fundamental magnitudes. Density, for example, is a derived magnitude, since it is measured as the ratio of mass and volume, which are both fundamental magnitudes. A fundamental magnitude can be measured without the previous measurement of any other magnitude; therefore, we will consider first the measurement of fundamental magnitudes. Weight and length would be two illustrations of fundamental magnitudes.

According to Campbell's (4) analysis, an extensive scale can be constructed only if an operation of 'addition' can be performed experimentally with the objects (or 'systems') being considered. Such an operation can always be found in the case of fundamental measurement *in Physics.*

For example, to 'add' lengths, two sticks are laid end to end; to 'add' weights, the two weights are placed in the same pan of a balance; to add electrical *resistances,* the two resistances are connected *in series;* to add electrical *conductances,* the two conductances are connected in *parallel;* and so on. However, Campbell maintains that when we examine the material of psychology no such operation has been found. For example, it is contended that there is no way to take the pitch of one note and add it to the pitch of another note, no way to add one brightness to another, etc. At least it is contended that as yet no way has been found. Besides defining some operation which may be designated addition, it is necessary to find *experimentally* that this operation of addition satisfies certain requirements.

First it is necessary to find experimentally that the sum of two magnitudes has a magnitude that is greater than that of either of the parts. This condition clearly obtains for the addition of length, weight, resistance, or conductance as indicated above. Density of liquids would be an illustration of a magnitude which does not satisfy this property. No way has been found of *experimentally* 'adding' two liquids, and obtaining a liquid of a density greater than either of the two components. Density is thus not a fundamental mag-

nitude. It can be measured only after other magnitudes have been measured. For the same reason, temperature is not a fundamental magnitude.

It is also necessary that the magnitude of a sum depend only on the magnitudes of the parts, and not upon the order or the method of addition (3, 4). It is conventional to state this general principle in three separate parts.

1. Equals plus equals gives equals.
2. Adding A to B gives a result identical with that obtained by adding B to A.
3. Adding C to (A + B) gives the same result as adding (B + C) to A.

These conditions are given as necessary for fundamental measurement (4, also 3, 6, 7, 8, 12, 17). It is contended, and no one has as yet successfully disputed this contention, that for the material of psychology, no operation of addition has been found that meets these requirements. Therefore it is argued that the psychological magnitudes of psychophysics cannot be fundamental magnitudes.

It is also asserted that the other types of magnitudes (derived and defined) depend upon the fundamental magnitudes, and are simply a function of one or more fundamental magnitudes. It is held that it is pointless to 'measure' sensations simply as some arbitrary function of the stimulus, so that sensations and other material of psychology cannot be meaningfully measured as derived or defined magnitudes. Therefore, since cessation magnitudes are neither fundamental, derived, nor defined magnitudes, they cannot be magnitudes at all (7, 8).

This line of argument demonstrates that it is possible to define 'measurement' in such a way that 'psychological measurement' does not satisfy the definition. A generalization of the "operation of addition which can be performed experimentally with the objects" will be presented. It will be shown that paired comparisons scales satisfy this broader definition, have useful properties not possessed by ordinal scales, and have, formally, a point-for-point correspondence with certain types of physical measurement.

No attempt will be made to decide whether scales constructed by the method of paired comparisons are fundamental, derived, defined, or some other type of scale. It will be shown, however, that these scales have not only all the properties of ordinal scales, but also many of the additional characteristics of 'extensive' or 'additive' scales used in physical measurements.

It does not seem that any special importance should attach to the operation of *addition*. The important thing is that numbers can be assigned to objects in such a way that an operation (not necessarily addition) with two numbers will indicate the result of some experiment which can be performed with the two objects represented by these numbers. This experiment should be *independent of the rest of the population of scaled objects*. In other words, a set of objects has been scaled when a number has been assigned to each object, and when some operation with these numbers enables one to predict the results of some experiment with the objects. This generalization clearly includes the case of addition given by Campbell. The *sum* of the two numbers representing two weights, gives the number of the weight that will balance those two weights. Also the *difference* between the two numbers representing any two weights is the number of the weight necessary to balance the scale when the two weights are on opposite sides. The same sort of illustrations could be given for length, electrical resistance, electrical conductance, etc. The operation of addition in these cases performed upon

numbers, gives another number, and this new number represents the results of an experiment that may be performed with the objects represented by the first two numbers. In the case of psychophysics, the operation is subtraction. In the case of fundamental physical magnitudes the operation is addition. It may be noted that for fundamental magnitudes, different investigators following the procedures outlined by Campbell would, when working in complete independence of each other, obtain scales which differed only in the arbitrarily selected 'unit of measurement.' Hence the scale developed by one investigator could be transformed into that of another by the equation $y = ax$. In the case of psychophysics the equation would be $y = ax + b$, since there is nothing in the method of psychophysics to insure that the value of b would be zero.[1]

IV. PAIRED COMPARISONS (ONE-DIMENSIONAL CASE)

It will be shown that scales constructed by the method of paired comparisons and Thurstone's law of comparative judgment (18) not only satisfy all the criteria of ordinal scales as indicated above, but also have the following characteristics:

1. Scale values can be assigned in such a way that an operation (subtraction) of any two scale values represents the results of an experiment with the two items designated by these scale values. Thus these scales illustrate a general characteristic of which the additive property of fundamental magnitudes is another special case.

2. The scale values are not dependent upon the particular population of objects chosen. This property removes another of the limitations of ordinal scales mentioned above. One of the properties of extensive magnitudes is that once the measurement system is set up, new objects will fit readily into the already established scheme. Such is also the case for scales constructed by the method of paired comparisons.

3. The scale values are not dependent upon any arbitrarily assumed relationship, or upon any experimentally determined relationship to another magnitude; hence such scales perhaps could not be regarded as either derived or defined magnitudes. However, psychological scales, as indicated later, have many properties of 'distance.' Since Campbell lists physical distance as a *defined* magnitude he might possibly feel that psychophysical scales are also illustrations of defined magnitudes.

In the method of paired comparisons the subject is presented with two stimuli and asked to judge which is greater with respect to some characteristic such as size, brightness, redness, beauty, etc. We can assign a set of numbers to these objects, such that the difference between each possible pair of numbers represents some experiment which can be undertaken with the corresponding objects, and which will satisfy the other requirements given above.

We can outline briefly the psychophysical procedure using the problem of scaling six lights with respect to brightness as an illustration. The lights are designated A, B, C, D, E, and F.

Step 1. Present the subject with each possible pair AB, AC, etc. to EF, and ask "Which is brighter?" Repeat this procedure for each of the fifteen possible pairs of lights until a reliable determination has been made of the 'percentage of judgments *i greater than j*' (where 'i' and 'j' assume in turn each value from A to F).

Step 2. Convert these percentages into 'distances' (from A to B, A to C,

[1] At least there is no such method that has been verified. One attempt (10) has been made to formulate a psychophysical method by which the zero point could be experimentally determined.

etc. to E to F) by some appropriate procedure. The appropriateness is to be checked in the next step. The procedure usually used has been to use tables of the normal curve, so that if the percentage of judgments "A greater than B" is indicated by the shaded area in Fig. 1, then the distance from A to B is the base-line distance indicated by the arrows.

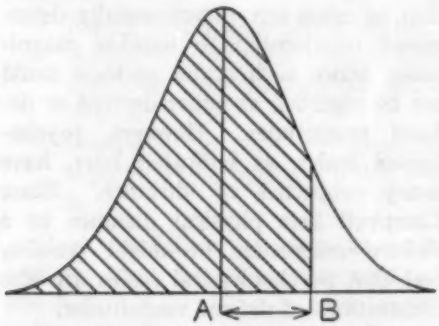


FIG. 1. The relationship between 'per cent of judgments A > B' and the 'distance between A and B' is shown.

Step 3. Check to see that these distances satisfy the criterion for a linear scale. In any linear scale the number representing the distance AB when added to the number representing the distance BC should give the number which represents the distance AC. This statement is true for all possible pairs of distances. The easiest systematic way to set up this check is to arrange these 'distances' as shown in table 1. In this table AB is used to represent the distance from A to B while BA (= − AB) represents the distance from B to A. The distance from an object to itself (AA, BB, etc.) is experimentally determinable if the objects cannot be identified uniquely. For instance, the brightness of two lights is the same when the percentage of judgments "i brighter than j" is 50, which indicates a distance of zero.

TABLE 1

THE DISTANCES 'ij' ARE SET UP SYSTEMATICALLY TO FACILITATE CHECKING FOR A LINEAR SCALE

	A	B	C	D	E	F
A	AA	AB	AC	AD	AE	AF
B	BA	BB	BC	BD	BE	BF
C	CA	CB	CC	CD	CE	CF
D	DA	DB	DC	DD	DE	DF
E	EA	EB	EC	ED	EE	EF
F	FA	FB	FC	FD	FE	FF

In actual experimentation it is usually assumed that AA, BB, etc. equals zero, and also assumed that ij equals negative ji. These assumptions can be checked experimentally.

The distances between adjacent points (indicated by underlining in the table) should, when added in various combinations, give each of the other entries in the table (except the principal diagonal entries of zero). For a rapid and systematic survey of the data, the easiest thing to do is to plot each possible pair of columns on a graph. The criterion for a linear scale is satisfied if each plot is a straight line with a slope of unity. It can readily be seen that there is nothing in the mechanics of the procedures outlined which forces these plots to all be straight lines with unit slopes, so that if this occurs it demonstrates a very interesting property of the judgments for that set of objects.

Step 4. If this internal consistency check is satisfied, it is possible to take the intercept of the plots of adjacent columns as representing the distances AB, BC, CD, DE, and EF, and to arbitrarily assign A the value zero.

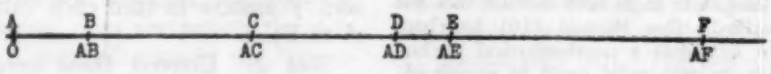


FIG. 2. One possible result of scaling procedures is illustrated.

Then B receives the value AB, C the value AB + BC, D the value AB + BC + CD, and F receives the value of the sum of all five distances (see Fig. 2). We have thus assigned a number to each of the six lights, and this set of numbers has a very interesting property. The difference between the numbers representing any two objects, such as i and j, will, with the aid of a table of the normal curve, give a close approximation to the percentage of judgments "i greater than j." That is to say, six numbers (one for each object) and one operation (subtraction) summarizes, or will enable one to predict, the results of $N(N-1)/2$ (in this case 15) independent experimental observations. The economy of representation is not great when dealing with only six objects, but in the case of 20 objects, for example, the same techniques, if successful, could be applied so that 20 numbers and one operation would summarize (or predict) the results of 190 observations which are experimentally independent of one another.

Let us see if this method satisfies the criteria which were attributed to it earlier.

1. Scale values are not dependent on the particular population of objects or systems chosen. It is clear that a row (and corresponding column) of table 1 may be removed without affecting the values remaining, and hence not affecting the scale values, unless the zero point happens to be the point removed, in which case there is a constant amount subtracted from each of the scale values. Also, if new points are added this does not disturb the already existing relations, and hence does not disturb any scale values unless, of course, one thereby chooses a new zero point.

2. The scale values are not dependent upon any arbitrarily assumed relationship to any other magnitude. It will be noted that the physical measurements of the stimuli need not be known at all, except that they might serve as convenient tags for duplicating the experiment. The criterion of internal consistency is judged to be satisfied, or not satisfied, depending wholly upon relationships within the dimension being considered.

3. The scale values have a meaning, in the sense that the difference between any two scale values for, say, i and j, indicates the percentage of judgments "i greater than j."

One way in which such a scale differs from fundamental magnitudes should be emphasized here. In the case of weight, for example, the two numbers (of weights) subtracted give a number which *also represents a weight*, namely, the weight it takes to balance the scales. In the case of psychophysics the new number is not the number of another sensation. It represents something quite different, namely, the percentage of judgments "$i > j$." This same property is possessed by many of the derived or defined scales, but is not a characteristic of 'fundamental magnitudes,' so that psychological measurement is not fundamental measurement according to such a criterion. The topic of defined magnitudes is discussed so briefly by Campbell that it is difficult to determine whether or not he would regard the scales of psychophysics as defined magnitudes.

It should be noted here that the foregoing illustration assumed only the simplest case of the law of comparative judgment, and assumed that the intermediate category was eliminated, *i.e.*, that all judges were forced to judge either that "i is brighter than j," or that "j is brighter than i." It is possible to handle the intermediate category, and it is possible to deal with cases other than the simplest one illustrated. However, the outline of the main points

is best presented if such complications are not introduced.

Also, it should be pointed out that several scaling studies have been successfully carried out using this simplest case; therefore, it is not so simplified as to be unusable. Thurstone (19) applied the method of paired comparisons and the law of comparative judgment to a list of 19 offenses, ranging from rape and homicide through perjury and smuggling to vagrancy. For each possible pair of offenses the judges were asked "Which is the more serious offense?" On the basis of the percentage of subjects judging "i more serious than j" these 19 offenses are each given psychological scale values such that the difference between these scale values represents the proportion of judgments "i more serious than j." That is to say, with 19 numbers and one operation of subtraction, it is possible to summarize (or to predict) the results of 171 operationally independent experiments $(18 \times 19)/2$. It should also be pointed out that we have here a case in which no measurement of the physical magnitude is possible, since there is no physical magnitude to correspond with the 'seriousness of an offense,' so it is clear that we are not dealing with a derived or defined magnitude in this case, unless the percentage of persons making a given judgment is regarded as the defining fundamental magnitude.

There is a similar application of the method of paired comparisons and the simplest case of the law of comparative judgment to the measurement of nationality preferences (20). In this study the stimuli used are the names of different nationalities, Norwegian, English, Turkish, etc. The judges are asked which nationality they prefer. In an unpublished master's thesis Frederick R. Eggan studied the effect of the form of the question upon judgments. For example, in constructing a scale of nationality preferences will one result be obtained when the question is, "Which nationality would you rather associate with?" and quite different results be obtained when the question is, "Which nationality would you rather your sister married?" Five different questions were asked of five different groups of subjects. One group was asked which nationality they would rather associate with; a second group, which they would prefer as a fellow student; a third group, which they would prefer as a neighbor; a fourth group, which they would prefer to have their sister marry; and a fifth group, with which they would prefer to do business. The rank order correlations among these five different sets of scale values ranged from .972 to .995. It is realized, of course, that as far as theory of scaling is concerned it is immaterial whether the foregoing experiment had revealed one scale for the five different forms of questions, or five scales, one for each form. However, if scale values for a set of nationality preferences remain the same with considerable variation in the form of question asked, one has a more general and hence a more useful scale than if the values varied with each change in the form of the question. For example, it is a very convenient property of length that it may be measured with a ruler, string, or by triangulation with the same results in all three cases. Furthermore, since the results in Eggan's experiment were the same for five different groups when each group was tried with a different form of the question, it is reasonable to suppose that if the form of the question had been the same, then the results would have been comparable. That is to say, the results are not dependent upon the group of subjects that do the scaling, provided, of course, that they do not have too widely different cultural backgrounds. For example, we would expect to find that the scale values of

'English' and 'Turk' found for a group of Englishmen would be different from the scale values found when the judges were a group of Turks.

The comparability of scales constructed by different psychological methods has also been studied (9). Again it is not logically essential that scales constructed by different methods should be linearly related to each other. However, one's confidence in his understanding of a scientific method is increased if it is possible to vary the method in accordance with certain principles and find that results are not affected. Hevner (9) found that the rank order method and the method of paired comparisons gave results that were linearly related to each other. However, this was not true for the method of equal appearing intervals when compared with either of the other two methods.

This problem was pursued further by Saffir (15). He substituted the method of successive intervals for the method of equal appearing intervals, and found linear relations between scales constructed by paired comparisons, by rank order and by successive intervals. It is interesting to note that Reese (12) criticizes Hevner, on the ground that there was no careful attempt to instruct the subjects to make the intervals equal.[2] When the possibility of unequal intervals which appears in the *experimental* conditions is allowed for in the scaling *theory* applied, then three different scaling methods all agree. Again it must be emphasized that while such agreement is not essential for the basic theory of scaling, it does increase one's confidence in the methods, and in the theory on which they are based.

[2] It is of interest to note that the discussions of mental measurement (1, 2, 7, 8, 12, 17) do not refer to *any* of the articles on psychophysics (9, 10, 11, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24) listed here, with the sole exception that Reese (12) refers to Hevner's study (9). In this citation, however, the significance of the agreement between the method of rank order and paired comparisons was completely ignored.

V. PAIRED COMPARISONS (MULTI-DIMENSIONAL CASE)

We have seen that there are several experimental studies which have successfully used the simplest case of the law of comparative judgment with omission of the intermediate category. However, in order to illustrate the scope of modern psychophysics we will consider one additional complication. What is the appropriate procedure when the check of internal consistency as listed in Step 3 above is not satisfied?

If the internal consistency check is not satisfied by any of the methods described by Thurstone (18), then there are two alternative procedures open. One is to seek some distribution law, other than the normal curve, such that when it is used the internal consistency check will be satisfied. The other is to assume that one is not dealing with a one-dimensional system and to see if the set of judgments will be 'consistent' if one adopts procedures proper for a space of dimensionality greater than one. Let us illustrate this extension in some detail for the case of six points in a two-dimensional continuum. In considering such problems as those cited above of nationality preferences or seriousness of offenses, it is not implausible to assume that in some cases the check of internal consistency for a linear continuum will not work.

It should be noted here that Campbell (3, 4) explicitly considers the case of two and higher dimensional continua, and says that it is impossible to regard objects scattered in such a space as being 'measured.' If the definition of addition as given by Campbell is rigidly adhered to, then it is probable that in a continuum greater than one-dimensional it would be impossible to define

a unique operation of addition. However, if one takes the broader definition previously suggested in Section III, then the difficulty disappears, since it is easy to have some operation that is performed with numbers in a two-space correspond to an experimental operation with the objects. Dr. Campbell points out that the objects in two-space can be represented by different coördinate systems, and hence there is no uniqueness in the representation by numbers. However, the mutual inter-distances of a set of points is the same regardless of the coördinates to which they are referred. If these distances have some experimental parallel, then the number system would be very useful. For example, points may be located on a map and by means of a coördinate system it is then possible to determine the distance between any two points on the map. To make a complete table of all the distances from each point to each other would be a much less convenient representation than would the map. In the case of psychological objects in a two-space (or higher space) the same thing is true. The interpoint distances are invariant, regardless of the coördinate system. Nevertheless, just as with the map, some coördinate system, even a completely arbitrary one, is more convenient than a table of all interpoint distances. Let us turn to a concrete illustration of this point.

Suppose we have the six points or objects shown in Fig. 3. For such a set of six points there are fifteen inter-point distances which are shown by the (solid and dotted) lines in Fig. 3. The points might, for example, be six offenses, as indicated above, to be judged with respect to relative seriousness, or they might be a set of six nations to be judged with regard to the 'probability of war' between each possible pair (11). The scaling procedure is to repeat Steps 1 and 2 as indicated above.

Step 1. Present the subject with each

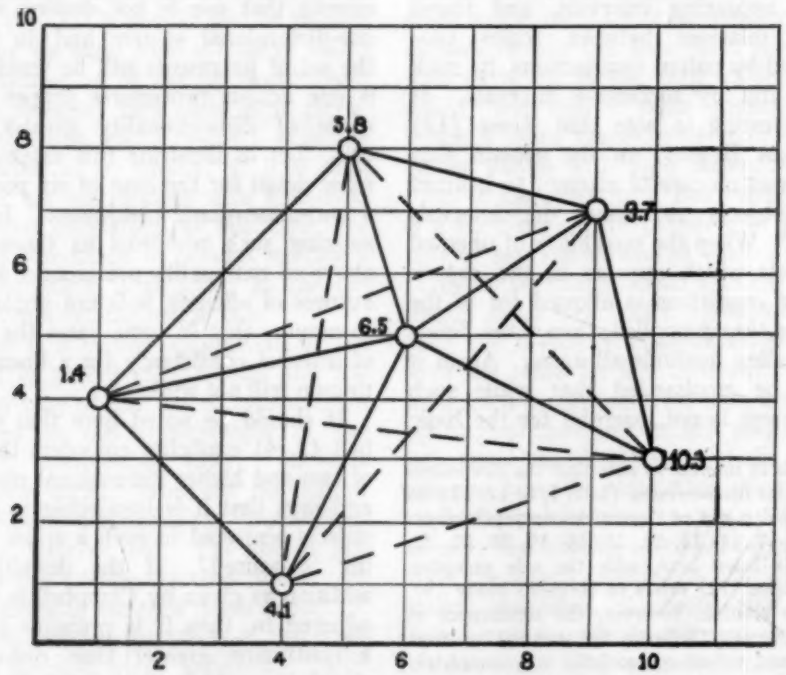


FIG. 3. The six points to be scaled are shown in a two-space.

possible pair of offenses and ask "Which is more serious?" Determine the percentage of judgments "i more serious than j."

Step 2. Convert the obtained percentages into distances by some appropriate procedure.

Step 3. Observe that there is *no* linear plot between all possible pairs of columns of the entries corresponding to those in Table 1.

Step 4. Since these inter-point distances uniquely determine the dimensionality of the space occupied by the set of points (23) it is possible to apply Young and Householder's test to determine the dimensionality of the space occupied by the set of points.

If our test shows that the points occupy a two-space, then an appropriately selected set of *nine* inter-point distances such as those shown by the solid lines in Fig. 3 will suffice to determine the other six distances. The converse also holds; namely, if a set of nine of these distances are necessary and sufficient to determine the other six, then the points are distributed in a two-space.

Step 5. It is then possible to assign coördinates to each point. Since we have a plane or a two-space, two numbers are assigned each point, as indicated in Fig. 3.

Step 6. By appropriate rules of vector subtraction it is now possible to calculate each of the 15 distances shown in Fig. 3 from the six coördinates (a pair of coördinates for each point) illustrated in Fig. 3.

We have here again a very useful numbering system, since from the six double coördinates and the appropriate operation of vector subtraction we can calculate each of the 15 experimental distances, or predict the percentage of persons who will judge one offense more serious than the other. New points may be added to the system without disturbing the numbering system of the old set. This is another way of saying that this type of scale has advantages beyond the ordinal scale. Likewise, the scale is completely self-contained with a two-dimensional, instead of a one-dimensional, check of internal consistency. It is not clearly a derived or defined scale, since it is not based upon observed or assumed relationships to other magnitudes. In other words, it has the major characteristics of the one-dimensional psychophysical scales discussed above.

It is to be noted that in dealing with a three-space or higher space we would proceed by the same steps, except that the coördinates assigned would be sets of three numbers to each point for a three-space, sets of four for a four-space, etc.

In each case a composite number would be assigned to each object, and these numbers together with an appropriate operation of vector subtraction would enable one to predict the results of a large number of experiments.

Again this multi-dimensional method has had a few experimental try-outs. It has been applied by Richardson (13) to a set of colors differing in gray-value and in saturation. A two-dimensional continuum was found adequate in this case. Klingberg (11) applied the method to a study of the likelihood of war between various nations.[3] He used a short scale separation to indicate a small probability of war and a larger one to indicate greater probability of war. He found that a three-dimensional continuum was satisfactory for representing the relative 'friendliness' of six nations,

[3] In these studies of Richardson and of Klingberg a slight variant of the method of paired comparisons called the 'method of triads' was used. Since the bases of the two methods are similar no detailed reconsideration is necessary. The method of triads uses paired comparisons of *distances* between objects, rather than of objects. Otherwise the considerations applying to one method also apply to the other.

Great Britain, France, the United States, Italy, Germany, and Japan. The inclusion of Russia added a fourth dimension to the system.

VI. PSYCHOPHYSICS COMPARED WITH SURVEYING

Let us recapitulate the main points of one-, two-, and higher dimensional psychophysics by showing the analogy between operations performed in psychophysics and those performed in surveying a tract of land. There is a close parallel between the operations performed in psychophysics and the procedure of a surveyor. If the surveyor is working in a *linear continuum*, that is to say, if he makes a map of a row of houses or other objects that lie in a straight line, he can measure the distance from A to B, from B to C, and from A to C, and then check his results by adding the first two distances to see if they equal the third one. Formally this is the same procedure that is followed in the linear case outlined above for the law of comparative judgment. However, the psychologist is more systematic and makes a complete set of measurements of distances between all possible pairs of points in order to get a complete set of checks. The measurements of the surveyor are more accurate, so he makes only a few of the possible checks.

Let us turn now to the two-dimensional case. If it is found that the distance AB plus the distance BC is *greater* than the distance AC (instead of being equal to AC) then one is dealing with a two-dimensional continuum or a plane surface. When the surveyor is working in a plane surface (as he usually is) arithmetic addition cannot be used to check his measurements. It is necessary to use a more complex formula to check his measurements. The psychologist would use the Young and Householder method for checking the dimensionality of a set of points. No one expects the distance from Chicago to Memphis plus the distance from Memphis to New York to equal the distance from Chicago to New York. We recognize here that we are dealing with a two-dimensional space, and instead of assigning each city one number to represent it, each location is assigned two numbers (such as latitude and longitude) to represent position in a two-space. As far as usefulness is concerned, for indicating locations on a map, any arbitrary set of coördinates would suffice. It happens that we have reached agreement on the particular set to be used. If the psychologist finds his material occupies a two-space it may be necessary to adopt an arbitrary set of coördinates until some agreement or method of reaching an agreement is decided upon.

As the surveyor extends the area surveyed he finds that the formulas suitable for a plane surface or a flat two-space no longer work. He is dealing with a curved two-space, the surface of the earth, and the distances will check only if a set of formulas is used which is suitable for such a space. So far in the psychological investigations no one has found a set of points, the inter-distances of which will check only if he uses formulas appropriate for a curved two-space.

If the surveyor is checking measurements between points in a mountainous terrain, then he must use formulas approriate to a three-space in order to check his results. In the psychological work the problem is similar. A three-dimensional space, and even a four-dimensional space, was found necessary for the results of Klingberg (11). To designate points in a three-dimensional space it is necessary to assign three numbers to each point. The surveyor would call these latitude, longitude, and altitude. Higher dimensional cases have not yet been sufficiently investigated in

psychology for any general agreement to have been reached on the particular axes to be used. However, any set of axes will specify the points so that the distances between them can be calculated from sets of three numbers, using the rule for vector subtraction.

It should be emphasized again that the method of paired comparisons and the law of comparative judgment deal solely with the psychological object. This object may be a sensation, which has a definite physical stimulus correlate, and the subject may be asked to judge which sound is louder or which is higher, or he may be asked to judge which light is brighter. In such cases the accuracy of the subject's judgment may be checked by physical measurement. However, the method of paired comparisons and the law of comparative judgment would apply just as well if the subjects were asked to judge which picture is more beautiful, or with which nationality they would rather associate. In these cases there is no physical correlate. There is no way to check the accuracy of a judgment. However, it is still possible to use these scaling methods and to assign to each object a number that is valuable, in that the difference between the numbers representing any two objects will indicate how the judgments will divide with respect to those two objects.

VII. SUMMARY

To summarize the line of argument presented here we may list the following major points.

1. The critiques of psychological measurement appearing during the past decade have carefully considered several different psychological scaling methods, but have ignored the significant features of the method of paired comparisons and the law of comparative judgment (18) as well as the various illustrations, developments, and extensions of this method appearing since Thurstone's 1927 article. It is necessary to consider this method carefully since it has several advantages over other psychological scaling methods.

2. This method satisfies all the requirements for, and properties of, ordinal scales.

3. In addition to satisfying criteria for ordinal scales, scales constructed by this method satisfy a generalization of the criterion specified by Campbell for additive scales. An operation (subtraction) with the numbers representing any two objects gives another number which predicts the distribution of judgments with respect to those two objects. This percentage can be verified in complete independence of the rest of the scale. This means that the simple device of using N numbers and one operation will predict (or summarize) the results of $(N-1)N/2$ independent experiments. However, it should be noted that we start with two numbers, each representing a *sensation*, and that the *difference* does not represent a *sensation* but a *percentage of judgments*. In the case of fundamental magnitudes, such as weight or length, the *numbers* representing two *weights* when subtracted give another number which also represents a weight. Scales in psychophysics do not possess this characteristic of fundamental magnitudes.

4. The scale values of a set of objects are not dependent upon the particular population of objects chosen. Once the measurement system is set up, new objects will fit readily into the already established scheme.

5. The scale values are not dependent upon any arbitrarily assumed relationship, or upon any experimentally determined relationship to another magnitude; hence it would seem that such scales cannot be regarded as either derived or defined magnitudes. (See 7 below.)

6. These methods have been tried out sufficiently in numerous experiments to indicate that the methods are feasible ones.

7. A close parallel can be traced between the measurement of distance in the physical sense and the scale values of psychophysics. Since distance is classified as a defined magnitude in Campbell's system (4, p. 95) it may be that the similarities between measurement of distance and the law of comparative judgment indicate that paired comparison scales should be classified as defined magnitudes. However, regardless of the details of subclassification of extensive magnitudes psychophysical scales have many of the important properties of extensive scales in physical measurement.

8. The method of paired comparisons has been extended so that objects may be scaled even though they are not in a linear continuum. Does such 'multi-dimensional scaling' constitute 'measurement'? Campbell (4, p. 12) maintains that multi-dimensional series probably do not form 'scales' unless they can be resolved into a combination of one-dimensional systems. In other words, Campbell *defines* measurement to include only linear continua. However, criterion 3 above, the generalization of Campbell's additive criterion, is satisfied for these multi-dimensional scales. Thus they do constitute 'measurement' if one accepts the generalization given in criterion 3 above.

It is hoped that this exposition of the characteristics of the method of paired comparisons and the law of comparative judgment will lead to a more critical examination of the properties of the method by persons interested in the logic of measurement. The developments of the last twenty years in this method furnish much valuable material which should not be completely ignored, as it has been in all recent discussions of the logic of measurement.

REFERENCES

1. BARTLETT, R. J. Measurement in psychology. Address to Section J, Psychology. *Report of the British Association for the Advancement of Science,* 1940, 1, 422–441.
2. BERGMANN, G., & SPENCE, K. W. The logic of psychophysical measurement. PSYCHOL. REV., 1944, 51, 1–24.
3. CAMPBELL, N. R. *Physics, the elements.* Cambridge: The University Press, 1920.
4. ——. *An account of the principles of measurement and calculation.* London: Longmans, Green, 1928.
5. ——. Symposium: measurement and its importance for philosophy. *Proc. Arist. Soc. Suppl.,* London: Harrison and Sons, 1938, 17, 121–142.
6. COHEN, M. R., & NAGEL, E. *An introduction to logic and scientific method.* New York: Harcourt Brace, 1934.
7. FERGUSON, A., MYERS, C. S., BARTLETT, R. J., *et al.* Quantitative estimation of sensory events. *Report of the Brit. Assoc. Advancement Science,* Cambridge, 1938, 108th year, pp. 277–334.
8. FERGUSON, A., MYERS, C. S., BARTLETT, R. J., *et al.* Quantitative estimation of sensory events. *Final Report, Adv. of Sci.,* 1940, 2, 331–349.
9. HEVNER, K. Comparative study of three psychophysical methods. Ph.D. dissertation, University of Chicago, 1928. Also published as, An empirical study of three psychophysical methods, *J. gen. Psychol.,* 1930, 4, 191–212.
10. HORST, A. P. A method for determining the absolute affective value of a series of stimulus situations. *J. educ. Psychol.,* 1932, 23, 418–440.
11. KLINGBERG, F. L. Studies in measurement of the relations among sovereign states. *Psychometrika,* 1941, 6, 335–352.
12. REESE, T. W. Application of the theory of physical measurement to the measurement of psychological magnitudes with three experimental examples. *Psychol. Monogr.,* 1943, 55, No. 3, whole number 251.
13. RICHARDSON, M. W. Multi-dimensional psychophysics. Paper read at meeting of American Psychological Association at Columbus, Ohio, Sept. 8, 1938, Abstract in *Psychol. Bull.,* 1938, 35, 659.
14. RUSSELL, J. T., & WRIGHT, Q. National attitudes on the far eastern controversy. *Amer. Pol. Sci. Rev.,* 1933, 27, 555–576.

15. SAFFIR, M. A. A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 1937, **2**, 179–198.
16. SCATES, D. E. The essential conditions of measurement. *Psychometrika*, 1937, **2**, 27–34.
17. THOMAS, L. G. Mental tests as instruments of science. *Psychol. Monogr.*, 1942, **54**, No. 3.
18. THURSTONE, L. L. A law of comparative judgment. PSYCHOL. REV., 1927, **34**, 273–286.
19. ——. Method of paired comparisons for social values. *J. abnorm. soc. Psychol.*, 1927, **21**, 384–400.
20. ——. An experimental study of nationality preferences. *J. gen. Psychol.*, 1928, **1**, 405–425.
21. ——. Theory of attitude measurement. PSYCHOL. REV., 1929, **36**, 222–241.
22. WRIGHT, Q., & NELSON, C. J. American attitudes toward Japan and China, 1937–38. *Publ. Opin. Quart.*, 1939, **3**, 46-62.
23. YOUNG, G., & HOUSEHOLDER, A. S. Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 1938, **3**, 19–22.
24. ————. A note on multi-dimensional psychophysical analysis. *Psychometrika*, 1941, **6**, 331–333.

REPETITION AND LEARNING. III. MEMORY AND THOUGHT

BY THOMAS W. COOK

Acadia University

I. Memory

The term memory has often been given a meaning wider than I intend to give it here. It has in fact been used to signify any effect of past experience on present performance. This makes memory synonymous with learning. There are, however, two features that might be thought to distinguish memory from other learning functions. These are reference to the past, and recurrence in the absence of the original stimuli. The former is usually called recognition, and the latter recall or reproduction. In the narrower sense of the term, we speak of recognition when we know we have had that experience before, and of recall or reproduction when the experience recurs without the original stimuli. But there is no such invariable association that would justify the application of a single term to phenomena bearing both of these and only these characteristics. Reference to a past is common, for example, in some forms of perceptual recognition with the original stimuli present, and, though it is not usually the case, recognition can accompany any sort of learned activity. On the other hand, occurrence in the absence of the stimulus can appear in its most complete form without past reference.

The relation of past reference to repetition poses some difficult problems. Elsewhere (4, 192) I have discussed some aspects of perceptual recognition, and as it is more convenient to study the role of repetition in recognition with the stimuli present, I shall not attempt to add anything here. This paper will be concerned with the most distinctive characteristic of the memorizing process, the faithful reproduction (within the desired limits) of the original experience in the absence of the stimuli.

(1) *Absence of the stimulus.* It is a curious paradox that in memorizing, where repetition is most prominent as a factor in learning, the final goal is to do without the stimuli. Repetition of stimuli helps us to become independent of those stimuli. Here in clearest form is the central difficulty of the relation between repetition and learning.

Behavioristic associationism has never accepted this interpretation, but has instead emphasized 'substitute' stimuli. That substitution of stimuli occurs is undoubted. In animal experiments its manipulation provides the chief control of the changes that occur during learning. And all human perceptions and skills are dependent upon stimuli in late as well as early stages of learning. In many cases the behavior acquires some relative independence of stimuli as learning progresses, but in the main the stimulus emphasis merely shifts. Even in memorizing or imagining there is an irreducible minimum of kinaesthetic and organic stimuli, and it is a reasonable guess that these are necessary for the occurrence of psychological processes. But it is conceivable that their role is chiefly maintenance of the framework or 'schema' within which learning occurs, and not the provision of substitute cues.

The test case that makes it difficult to accept the explanation of human memory in terms of substitute stimuli is the retention of a sensory pattern after one presentation of the stimulus. I sit at

the window and glance over the landscape. Then I shut my eyes and attempt to recall the picture. Before me is a lawn crossed by a road, in the left foreground a telephone pole, further back a house, and to the right another long building. I open my eyes and realize that I missed a lot, but how did I remember so much from one stimulus presentation? It seems improbable that the substitution of organic or kinaesthetic stimuli had anything to do with it other than provide the experience of my body as a frame of reference.

It would be a mistake, however, to regard ability to retain and report memory 'images' as independent of repetition. For this ability depends in turn upon the earlier development of special attitudes, and in all probability upon language. And these are not learned without repetition. On the other hand, the capacity upon which the report of imagery is based must in turn be grounded in perceptual organization and in the retention of such organization upon the removal of stimuli, since it appears in fairly complete form in young children. Nor is it primarily dependent upon overt motor expression, since the experience is in the psychological field whether I report it or not.

(2) *Organization of material beyond the memory span.* Because of its usefulness as an aid to recall when the original stimuli are no longer present, the development of a system of reference that will organize past experience and make it available has been highly motivated. The record of any event can of course be preserved through language. But to place that event properly in the past a system of reference is necessary. All of us make some use of the fact that there is a tendency for the sequence of remembered experiences to follow the original sequence of events. In addition we have individual reference systems based on the temporal order of key events in our lives. But for the most part we depend upon the clock and the calendar, which use the order of the number system to signify temporal sequence.

Most of the laboratory work on memorizing has been done with verbal materials. From the memory point of view these are usually classed as rote and logical, or nonsense and meaningful. In many ways meaningful material is more easily remembered than is nonsense material. Its memory span is wider, amounts beyond the span can be mastered in fewer repetitions, and the curve of forgetting drops more slowly. Moreover, English, Welborn, and Killian have shown that the same passage may be divided into rote and logical (v and s) items, of which the v items are forgotten more quickly than the s items (5).

Though this is not the place to attempt a complete analysis of so complex a problem, several aspects of the dependence of memory on meaning stand in such obvious relations to repetition that they should not be overlooked.

(*a*) There is at least a partial truth in the common assumption that the superiority of meaningful material is due to familiarity or repetition, and that meaningful material is only more completely memorized nonsense material. Certainly 123456789, occurring in a series of otherwise nonsense numbers, are more easily remembered than is the remainder of the series.

(*b*) The part played by similarity or repetition in the organization of art forms is a worth-while source of material for the psychologist. Since most of the laboratory work on memorizing has been done with serial materials, I shall discuss only the temporal arts. For analytic purposes music is the best example, since the memory value of its patterns is relatively uncomplicated by the presence of other types of meaningful organization. But for illustrative purposes poetry is more convenient,

since its repetitive devices have been to a greater extent analyzed and given names such as rhyme, metre, alliteration, and assonance. Now, there is little doubt that this type of repetition aids remembering.[1] But, besides prose sense, there are two factors which make experimental analysis unusually difficult. In the first place, poetic or musical patterns are not just a matter of similarity, but of more or less regular recurrence of similarity in difference. Within a single line of poetry there are many varieties of recurring pattern, and the problem of assessing the memory value of each is complicated not only by the interaction of rhyme, metre, alliteration, and assonance, but by breaks in the regularity.

In the second place, the aim of the poet or composer is not primarily to make something easy to remember. He uses devices that have high memory value in order to effect a more or less unitary impression on his audience, and he can only do this if we remember what went before. But his chief object is to induce a certain kind of experience. And for this purpose change is as important as constancy, for surprise and disappointment as well as fulfilment and satisfaction are integral features of the pattern. It is possible that for a person sufficiently alive to the delight of these surprises they have high memory value. On the other hand, I have gathered some unpublished data which indicate that the average reader finds unusual twists more difficult to recall than the remainder of the passage.

(*c*) A third aspect of similarity in memorizing is derived from its relation to generalization.[2] The emergence of generalizations also presents problems in the relation of repetition to thinking. But with these we are not here concerned, only with the way in which generalizations once made help us to remember. This is achieved to a degree by letting a sign (word) stand for a similarity. We are then able to remember a large number of objects almost as easily as one. This only holds true, however, if we are concerned exclusively with those features of the objects covered by the generalization. Thus if we have 1000 nonsense syllables we can bring these within the memory span by the phrase '1000 nonsense syllables,' and we can specify further general aspects by the words 'in serial order,' 'black,' and so on. And we could if required produce a list according to these specifications. But the concrete, individual features of the nonsense syllables, and their correct order, we could not reproduce without many hours of memorizing.

At first thought it seems plausible that there might exist systems of generalizations so logically related that all concrete details are carried by the pattern, and no rote memorization is necessary in order to remember it completely. Scientific and mathematical systems most closely approximate this situation. But no complex system can be mastered without language, and no language can be mastered without a minimum of repetitive memorizing. The problem is also complicated by the fact that discovery of the essential relations in a complex field can proceed only step by step, so that before a learner can master the more intricate relations he must already have a well-organized body of knowledge in the field. Without such a possession the materials suitable for generalization (if they are differentiated at all) are a mass of similar interfering elements.

(3) *Rote memorizing.* We are told by Gestalt psychologists that rote learn-

[1] Bell (1) found stanzas of poetry better retained than were prose passages made from the same words in slightly different order.

[2] Not the C-R variety. The present usage follows Humphrey (10).

ing is a form of organizing. This may be conceded to the extent that learning will not occur unless the disconnected elements are unified by a set or attitude, and that the finished product has many properties of organized processes. But when this is granted, it still remains that the action of repetition or similarity in rote memorizing has many distinctive features.

A useful starting point for discussion is Koffka's hypothesis that repetition causes the formation of aggregated trace systems which (*a*) have as many members as there were repetitions, (*b*) facilitate the memory of the trace system as such, and (*c*) inhibit the memories of the individual repetitions, just as aggregation of like elements in a series is supposed to inhibit memory of their individual peculiarities. By this hypothesis he explains the difficulty of learning nonsense syllables, retroactive inhibition, and other well-known phenomena of rote learning (13, 545). Though I believe that this theory represents a great advance in our understanding of the problem, it seems to me at best badly stated, and at worst to contain an error in formulation that might retard progress. My chief difficulty is that the phrase 'aggregated trace system' does not seem to make sense. The terms 'aggregation' and 'trace system' appear to be contradictory. A trace becomes more systematized by becoming better articulated, and articulation and aggregation surely have little in common. My own interpretation of the experimental data is that repetition increases articulation in one dimension and aggregation in a different dimension.

Perhaps I can avoid wrongly attributing to Koffka my changes and additions to his theory (since my debt to him needs no emphasis) if I restate the problem and its solution in my own words.

(*a*) My formulation starts from the fact that a series has both unity and variety. It is therefore reasonable to suppose that the formation of its trace has two corresponding dimensions, organizing the trace into a unity, and accentuating the differences within the trace. The former is usually called the formation of bonds. The latter is commonly overlooked, but since the stimulus-conditioned process has both unity and variety, it is as reasonable to suppose that removal of the stimulus weakens the differences between the members as that it diminishes the unity of the total trace. Repetition might then not only tie the members of the trace into a pattern, but gradually bring the differences between its members to the level of the stimulus-conditioned state.

(*b*) There are differences in the stimulus-situation other than those between members of the series, and these differences also may be weaker in the trace than in the stimulus-conditioned process. The differences between the series and its background we may pass over, since these are usually so striking that in all probability the figure-ground differentiation is well represented in the trace. It would, however, be easy to diminish figure-ground differences to such a degree that they would offer difficulty in learning. More pertinent are the differences between the trace of one repetition and that of the next. That the two must be to some extent differentiated in the trace system is deduced by Koffka, on the ground that we can remember two or more repetitions as distinct events (13, 440). It should be noted, however, that in serial learning these differences are not nearly so great as the differences within a series. The stimulus-conditioned difference between two repetitions is only the occurrence of no stimulation or of background stimulation. On the other hand, the members of a series (of digits, for example) are differentiated from one another not only by an interval between them, but by dif-

ferences of form between the members of the series. In order for its units to be as much alike as are successive repetitions, a series would have to consist of identical members, as ------ or 222222.

(*c*) In repeating a series mental set favors both unity and diversity within the series. Inter-repetition differences are favored by no such set, for increase of unity between the members of a series of repetitions is so valueless that no one would think of attempting to acquire it. Ordinarily if we want to remember the individual repetitions it is much more convenient merely to count them. It should be noted, however, that achieving unity-in-diversity within a number of successive repetitions is not an impossible task. One can take any number of identical repetitions (say 20) as a series. Eventually if we go over these 20 repetitions (no more and no less) again and again for a long enough period, we will finally be able to remember them in another fashion than that of designating them by a number. The series of 20 repetitions will gradually become an organized trace, unified but with each member distinct.

In view of the difficulty of the problem, perhaps it would be worth while to put the situation more concretely. Suppose we consider three repetitions of a series ABCDE.

Repetition 1	A B C D E
Repetition 2	A B C D E
Repetition 3	A B C D E

Now, supposing that after repetition 1 has ended, the trace of A B C D E is less unified and differentiated than it was with the stimulus present. When repetition 2 occurs, it increases both the unity and diversity of the trace of repetition 1. This is 'horizontal organization.' At the same time, the difference in time between repetition 1 and repetition 2 causes vertical organization of the series, much weaker than the horizontal organization, both because the inter-repetition differences are much smaller than the inter-series differences, and because mental set favors the horizontal dimension. Repetition 3 increases the horizontal unity and differentiation, and decreases the vertical differentiation, and so on.

If now we leave the consideration of complete identity, and look at the effects of partial identity, it will be found possible to formulate a plausible quantitative statement of the effects of inter-series and inter-repetition differences. On the one hand, the greater the likeness between A B C D E and its successor, the more will the second series benefit by the first. This is the well-known type of positive transfer usually called practice effect. For example, memorizing one nonsense syllable list helps most in memorizing another list of nonsense syllables, less in subsequent memorizing of digits, and still less in mastering more widely dissimilar materials such as prose or poetry.[8] On the other hand, the greater the differences between A B C D E and its successor, the easier it is to remember them both as distinct events. It seems quite likely that Koffka has here hit upon some reciprocal function with a common base—possibly that a memorizing *process* is favorably affected by a trace to the degree that they blend, and conversely, to the same degree the *traces* of the two are subsequently difficult to hold apart as separate events.

So far we have been discussing the effect of the trace of (*a*) the first repetition of a series, upon (*b*) the later repetitions of the same series, and also upon

[8] Both a-priori considerations and the experimental results suggest that on the average positive and negative transfers of the paired associates type (usually called new S to old R and old R to new S) are approximately equal and balance each other, and that the predominantly positive transfer is due to other factors.

the tendency for the traces of (*a*) and (*b*) to remain distinct. But when we consider the 'retroactive' function of repetition, the data require a different formulation. When two repetitions are identical, there is reason to believe that the influence of (*a*) on (*b*) is the same as the influence of (*b*) on (*a*), and that both are favorable. But the effects of partial identity of (*a*) and (*b*) upon the recall of (*a*) present a different picture. This is the familiar retroactive inhibition set-up. Here the evidence suggests that the true answer is along the general lines postulated by Gibson (6). Identity aids recall, because between identical repetitions there is no need for differentiation. But when we get less than complete identity, the greatest amount of retroactive inhibition is produced by material of lowest discriminability.[4] This would mean that interference of a process with a trace, and (or) interference between two similar traces at later recall, will decrease as ease of discrimination between the two increases.[5] This hypothesis fits all the data of McGeoch and his associates (14, p. 511). It may also bring into line the apparently variant results of the Robinson-Harden-Kennelly experiments (16, 7, 12). These could result from a combination of the two types of effect mentioned above. (*a*) When ABCD is presented, followed by different numbers of new and old letters AFCD, AFCX, EFCX, EFKX, the 'identical' letters may be sufficiently segregated to act like identical repetitions, and facilitate recall of the same letters in the preceding series.[6] (*b*) The inhibiting effect of the new letters on the recall of ABCD is a function of their discriminability. Since in the order given above there are 1, 2, 3, 4 letters to discriminate from ABCD, the difficulty of discrimination should be least with AFCD and increase with the number of letters to be discriminated, being greatest with EFKX. EFKX should thus have the greatest inhibitory effect on the recall of ABCD, and AFCD the least. Both applications of my adaptation of Gibson's hypothesis could probably be tested by experimental analysis along the lines suggested by Boring (2).

Although limitations of space prevent it here, there should be no difficulty in making a preliminary test of the adequacy of these hypotheses by attempting to fit them to the data accumulated in other experiments on recall. For example, the hypotheses fit Von Restoroff's results on crowded and isolated materials, since the better remembered isolated materials are also more discriminable. In this connection there is a limiting factor that should be mentioned: the differences in question, in order to effect inhibition, probably must be demanded by the conditions of reproduction. For example, Pillsbury and Rausch found that differences of color within a series would not produce the Kohler-Restoroff phenomenon (15). And I suspect that differences of color between two visually presented repetitions of a list of nonsense syllables also would not affect the later recall of either, since we are accustomed to neglect color both in learning and reproduction.

[4] A deductive system worked out on the basis of my assumptions would of course differ in many points from Gibson's. But much of our apparent disagreement is due to differences in terminology, and a great deal of the remainder could be arbitrated by experiment.

[5] For simplicity's sake I have omitted discussion of proactive inhibition.

[6] Boring (2) is probably correct in saying that we should count numbers of new and old letter combinations rather than individual letters. But the above account is easier to follow, and once understood can be translated into Boring's terms.

II. THOUGHT

The term thinking has several connotations in psychology. As related to learning, however, it is a form of problem-solving using imagery and (or) language. In its most characteristic form thinking is non-perceptual. Of course only visual and auditory stimuli are wholly missing. Kinaesthetic stimuli at least remain, but so far as thinking is symbolic this is of secondary significance. Any sign can carry any meaning, and since the manipulation of meanings is crucial in thinking, whether a particular meaning is carried by stimulus-conditioned or imaginal processes is of secondary importance. It may be merely an accident of development.

(1) *The material of thought.* Humphrey classifies the problems in this field into those concerned with the material of thought, and the arrangement of thought (9). The first of these we have already met with in memory and perception. Now, it is often assumed that the materials of both memory and thought come from past experience. We have seen, however, that memory is not reproductive in any absolute sense. In memorizing we learn to do without perceptual aids, and in any specific act of memorizing, we select only those materials or aspects of the material that we want to learn. Not one subject in a dozen remembers the color, size, or style of type of a series of nonsense syllables, only the distinctive form of each syllable and their serial order. On the other hand, however, memory is supposed to be repetitive in the sense that some features of the original experience are faithfully reproduced. A good memory is one that accurately retains and reproduces those features. Thought, on the contrary, does not aim at reproduction. Its goal is the solution of a problem, and it would be surprising if memory materials were not changed in the process.

In America at least the experimental study of thinking did not get well under way until after the Titchenerians had decided that imagery is the only reproduced mental content. Memory images were thought to be copies of sensations, and were classified in the same categories as sensations, visual, auditory, etc. The Würzburg experiments, however, called in question the dominance of imagery in thinking. The resulting controversy centered on two issues: whether a new material of thought had been discovered and whether thinking can proceed without images. The failure of either party to convince the other on these two points tended to obscure the positive results of the experimental work, namely, that there are great differences in amount, modality, and vividness of imagery used in thinking, that these differences are unimportant in the solution of many thought problems, and that in thinking on abstract topics the imaginal content may shrink to a near minimum. We should not, however, conclude that imagery is unimportant in thinking. While its exact role has not been definitely decided, there is no doubt that many problems are more easily solved by its aid.

That thinking can be relatively independent of any specific modality, quality, or intensity of imagery might be decided from other evidence. In an earlier paper I have suggested the probable basis of perceptual sign-learning (4). Certain perceived movements or other sensory patterns become signs of something about to happen. Now, these signs are mediated by objects and can have a full sensory content. But this is only because in ordinary experience the stimulus properties vary within a narrow range. Actually only a fraction of the stimulus is needed to arouse the perceptual sign, and only a fraction of the perception is needed to serve as a sign of coming events. All that is required

is that the sign be readily discriminable from other signs of different events. The resulting tendency to simplification of signs appears in marked form in written language, where it is only kept within bounds by the need of moderately fixed conventions for communication, and more recently by standardization for mechanical recording. In thought there are no comparable checks, and it would not be surprising to find that the sensory content of thought reduces to the lowest terms compatible with its sign function. This is indeed definitely true of my own thinking. I can on occasion get fairly vivid and detailed imagery. But my usual thought processes are governed by something like a 'law of least sensory content.' I tend to use as little imagery as possible for the purpose in hand.

Besides imagery, three chief types of material have been postulated, relations, attitudes, and whole qualities. Introspective study of relations during thinking is very difficult, but if we ask whether an intelligent adult solves problems *as if* he grasped relations, there can only be an affirmative answer. Some of the most dramatic developments in this field have come from Gestalt experiments with children and animals. But adult human problem-solving via language still presents the most unambiguous evidence. This is currently recognized in the statement that human adults abstract, generalize, and form concepts. What is seldom realized, however, is that the evidence is cumulative. Relations form systems of which the higher order relations depend upon the lower order relations. And it is easy to demonstrate, introspectively as well as by functional analysis of behavior, that a subject cannot master the higher order relations unless he already knows the crucial lower order relations.

The 'sets' and 'determining tendencies' of the Würzburg group depend chiefly upon functional analysis for proof of their existence. As Humphrey has pointed out (9), attitudes were conceived by their discoverers as a new type of content to be added to the imaginal elements of the prevailing sensationism. However, the changed emphasis in Gestalt psychology does not mean that they have no systematic use for the concept. Koffka retains attitudes as defined by the Ego-Object relationship. Gestalt psychology has indeed been more interested in dynamics than in classification. But when their descriptive terms are lifted from the context of specific experiments, what amounts to a novel class of content emerges. Thus the most obvious products of sensory organization, which Koffka calls 'things,' and 'framework,' are neither sensations, images, relations, nor attitudes. And in his attack upon *Organizing and Memorizing* by Gestalt techniques, Katona finds evidence for a novel class of content which he calls 'whole qualities' (11).

Katona's attempts to relate his findings to the concept of repetition indeed bring out clearly the difficulty of fitting this concept into a dynamic psychology. Freud, for example, hardly uses the concept of repetition. However, if repetition is to be fitted into a dynamic context it will not do to say vaguely that a pattern or process is repeated. One must discover what constant and what variable features are essential to a particular learning task.

How, then, can the material of thought be related to the idea of repetition? As far as concerns the relations between earlier and later phases of a particular reasoning process, Humphrey's notion of a four dimensional situation is more applicable than repetition (8), since a given reasoning process is psychologically continuous. Such breaks as occur do not mark off psychological realities. Thus if we decide that goal-set is a necessary feature of reason-

ing, it is perhaps more accurate to say that it is present continuously than that it is repeated. Even an enforced break through interruption or fatigue does not make the next trial a repetition. The tension remains, since the problem is unsolved, and once the problem *is* solved later trials become exercises in memorization. Yet in my view certain aspects of the problem of repetition are closely related to this constancy of content within a given act of problem solving. The issue in both cases has this common feature: what constant and what variable temporal features are most favorable to a solution?

(2) *The order and succession of thoughts.* Two types of observational fact have furthered the associationistic tendency to conceive novelty in thinking as change in the order and arrangement of material rather than in the material itself. (*a*) It is convenient in experimental work to arrange objects in the environment in such a way that the solution of a problem depends upon their rearrangement, either physical or ideational. (*b*) Language, the tool of thinking, seemingly consists of a number of discrete words, and a given thought problem is solved when we can rearrange these in an appropriate pattern. But the fixity and discreteness of words are illusory. As signs they might have had any of a wide variety of patterns. The distinctive novelty of thought cannot therefore lie in their rearrangement. On the other hand, when viewed as meanings and relationships they are seen to vary with learning and with context. Even moderately fixed meanings apply only to sentences or paragraphs. The word 'man' may vary from extreme particularity, as 'that man,' to the broad generality of man as a species. Therefore, while study of language is valuable, even essential, we shall discover the role of repetition in thinking only by going behind the words (as sensory forms) to their meanings, and tracing these through successive phases of problem solving.

One other traditional aspect of the change-constancy problem in thinking should perhaps be mentioned. During the pre-solution period, what parts, phases, or lines of attack are taken up, laid down, and taken up again? The names usually given to this characteristic of reasoning (multiple response, varied attack, trial and error), emphasize the change. But repetition as well as change occurs, and it is interesting to speculate on its relation to the efficiency of thinking. Here we meet a certain opposition between the two. Spence, indeed, in a different context, cites this opposition as a logical flaw in Krechevsky's concept of 'hypotheses' (17). Spence's argument, slightly paraphrased, runs as follows:

1. If a hypothesis is correct, it leads to a solution; that is, it is not repeated.
2. But we do not call it a hypothesis unless it is repeated; that is, unless it is wrong.
3. But to repeat wrong behavior is stupid.
4. Therefore, hypotheses are stupid.

This argument is irrefutable if one grants the implicit assumption that hypotheses are units of an all or nothing character. But if hypotheses are dynamic patterns that may be partly wrong and partly right, Spence's conclusions do not follow. For a hypothesis is stupid only if one contrasts it with a perfect immediate solution. In comparison with random behavior it may be brilliant.

Spence's dilemma might, however, be taken to indicate that with order as with material, a dynamic interpretation of thinking has little place for 'repetitions' as distinct events. For the dynamics of thought is most in evidence when it is

most efficient, and to the degree that it is efficient, thinking is free from error. If all the requisite data are presented simultaneously, a solution wholly determined by the internal structure of the system should appear at once. If the data are presented successively, the process will require time, but need involve no different principles from those that explain the immediate flash of insight. To the associationist such a solution is merely a special case of reproductive memory. Associationism thus accounts for thought by virtually denying its occurrence except in conjunction with error. By the Gestalt psychologist errorless thinking is ascribed to 'closure,' 'the law of fittingness,' or 'determination by internal relations.' In terms more directly borrowed from physics, a thought process is frequently described as a system under stress, and a solution as reduction of the tension. But if the physical analogy has more than suggestive significance, it represents a hope that both constant and repetitive features of problem-solving will eventually be formulated in 'laws,' from which can be derived not only the properties of static systems, but also the directions, kinds, and rates of change that occur when the equilibrium is disturbed.

Accordingly, I interpret Spence's dilemma to mean (in Gestalt terms) that error is the introduction of forces not congruent with those internal relations or tensions, which in the absence of such intruding factors would effect a solution. Such intrusions might be due to the large number of elements involved, to inadequate memorization of part aspects, to erroreous 'generalizations,' to 'fixations,' etc. Any repetitive trends occurring in successive trials (hypotheses) might then be a resultant of factors favoring a solution, and of factors working against a solution.

There remains, however, one aspect of thinking for which even a perfect immediate solution seems to demand repetition. This is the process of generalization. Since the abstract and general aspects of experience retain their most essential characteristics whether the stimuli are present or absent, they form a major link between thinking and perceptual problem-solving. Most of the experiments on generalization have been performed with the stimuli present, and it has often been tacitly assumed that stimuli provide the material for the generalizations. To my view, however, stimulus repetition only affects generalization indirectly through conditioning the perceptual field. The identities in question are parts of the behavioral environment, not the stimulus. Normally the context is also variable, so that the identities have to be discovered (abstracted) before they can be generalized. When context variability is at a minimum, generalization is easy. No one has any difficulty in cognizing that the dots in the series . . . . . are identical. The only aspect that must be 'abstracted from' is their relative position. What makes the process of abstractive generalization difficult is variation in kind and amount of context, plus the fact that we are not usually concerned with finding *any* generalization, but with discovering the generalization that will solve the problem in hand, a process that involves the relevance of the generalization to the total situation. The study of the forms and degrees of such variation in context upon the processes of abstraction and generalization is indeed one of the most important aspects of the change-constancy problem in thinking. However, since most of the previous work has been done with the stimuli present, and since I have nothing to add to Humphrey's careful analysis (10), I shall not further discuss the problem here.

In review, we may summarize the main features of this discussion of thinking:

(1) Memorizing aims at retention and reproduction of selected aspects of the psychological field, without the original stimuli. Thinking involves other varieties both of constancy and change, in that its end-result is to some degree novel, and that in many cases trial and error introduces repetition during the pre-solution period.

(2) The change-constancy problem in abstraction and generalization has many features common to perception and thought. The major difficulty is not in noting an obvious identity and giving it a name, but in discovering novel and relevant identity-difference patterns in complex situations. This in turn depends upon the acquisition of an interrelated and hierarchal system of references with appropriate language signs.

(3) Problems of thinking have been divided into those concerned with material and with order. Materials postulated have been images, relations, attitudes, and whole qualities. Images were originally conceived as reproductions of sensations. In thinking, however, sensory modality and quality chiefly function as carriers of meaning, and need retain only enough of their perceptual properties to permit their discrimination from one another. Abstracted and generalized aspects of experience, on the other hand, are not readily fitted into the image category, and when thought of as meanings and relations problems of material overlap with those of order, since in this case the relations are the materials, and at the same time order is itself a relation.

(4) The Gestalt view of problem solving implies that thought is errorless to the extent that it is determined by an interdependent system of internal relations. The expectation is that laws governing the behavior of such systems can be formulated (like their physical analogies) to cover such repetitive features as constancy of rate of change in a given direction. Errors would then be attributed to the intrusion of forces extraneous to the system, and repetitive trends such as 'hypotheses' would be resultants of the interaction of internal forces favoring a solution, and of external intrusions.

REFERENCES

1. BELL, R. Unpublished M.A. Thesis. Univ. of Toronto, 1945.
2. BORING, E. G. Communality in relation to practice and retroaction. *Amer. J. Psychol.*, 1941, **54**, 280–283.
3. COOK, T. W. Repetition and learning. I. Stimulus and response. PSYCHOL. REV., 1944, **51**, 25–36.
4. ——. Repetition and learning. II. Perception. PSYCHOL. REV., 1945, **53**, 187–198.
5. ENGLISH, H. B., WELBORN, E. L., & KILLIAN, C. D. Studies in substance memorization. *J. gen. Psychol.*, 1934, **11**, 233–259.
6. GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. PSYCHOL. REV., 1940, **47**, 196–229.
7. HARDEN, L. M. A quantitative study of the similarity factor in retroactive inhibition. *J. gen. Psychol.*, 1929, **2**, 421–430.
8. HUMPHREY, G. *The nature of learning*. London: Kegan Paul, 1933.
9. ——. The problem of the direction of thought. *Brit. J. Psychol.*, 1940, **30**, 183–196.
10. ——. The problem of generalization. *Bull. Can. Psychol. Assn.*, 1944, **4**, 37–50.
11. KATONA, G. *Organizing and memorizing*. New York: Columbia Univ. Press, 1940.
12. KENNELLY, T. W. The role of similarity in retroactive inhibition. *Arch. Psychol. N. Y.*, 1941, **37**, No. 260.
13. KOFFKA, K. *Principles of Gestalt psychology*. New York: Harcourt, 1935.
14. MCGEOCH, J. A. *The psychology of human learning*. New York: Longmans, 1942.
15. PILLSBURY, W. B., & RAUSCH, H. An extension of the Kohler-Restoroff inhibition phenomena. *Amer. J. Psychol.*, 1943, **56**, 293–298.
16. ROBINSON, E. S. The similarity factor in retroaction. *Amer. J. Psychol.*, 1927, **39**, 297–312.
17. SPENCE, K. W. Continuous vs. non-continuous interpretations of discrimination learning. PSYCHOL. REV., 1940, **47**, 287.

THE HISTORY OF DREAM THEORY

BY HAROLD GRIER McCURDY

Meredith College

The handling of the ancient history of dream theory in Freud's *Interpretation of Dreams* piques the reader's curiosity. The references are sketchy, they are apparently from secondary sources, and they omit a number of important names. The aim of the present article is to supplement the historical section of Freud's work by a more detailed examination into some of the ancient writers. Unfortunately, Büchsenschütz's essay on which Freud relies for information about this period is unavailable to me.

I

Three essays by Aristotle in the *Parva Naturalia* are the chief contribution on dreaming preserved to us from classical antiquity.[1] Taken together with a few remarks from Plato, and with due allowance for the partially obsolete physiology, they provide a theory which is decidedly modern in tone.

The key concept for Aristotle is sleep. Sleep opens the way for dreaming by reducing the intensity of sense-activity. This torpor of the senses is an expression of a change in the central organ of consciousness, the heart, which, because it has become deeply involved in the physiological processes of nutrition accounting for the state of sleep, no longer functions at maximum efficiency. The organs of sensation, being subsidiary to the heart, are likewise affected. There is thus an inverse relationship between nutrition and sensation. When nutritive activity dominates the organism to a very high degree, as in the sleep of early childhood, sensation drops to an approximate zero; it is for this reason that young children do not dream. Dreaming is a low-level sensory activity; it cannot occur when the heart and sense organs are completely absorbed in nutrition. But neither can it occur when the senses are fully active. Sleep is a prime requisite.

But what is the nature of sensory activity during sleep? Aristotle approaches this question after taking note of certain waking phenomena. He demonstrates by reference to after-images and sensory adaptation that the activity of the sense organs is normally prolonged beyond the moment of actual perception; or, in other words, that stimulating objects set up processes in the sense organs that continue after the objects are withdrawn. He demonstrates besides, by an example which sounds curiously magical but which may not be entirely so, that the sense organ may impose some of its own qualities on the object.[2] The senses, then, carry on certain activities which are not strictly dependent upon the external stimulating objects. These activities, or the images to which they give rise, may be mistaken for real objects, especially when emotion is present. He supports this conclusion by several examples. He then touches on the important idea that judgment regarding perceptions or images and the power of having them (phantasy) are separate functions. It follows logically that consciousness may

[1] 'De somno et vigilia,' 'De somniis,' 'De divinatione per somnum.' In *The works of Aristotle* (W. D. Ross, Ed.; trans. by J. I. Beare). Oxford: Oxford Univ. Press, 1931, Vol. 3.

[2] *Ibid.*, 459b. "The gaze of a menstruous woman tarnishes mirrors with a bloody mist." Meant as a subjective effect, this phenomenon might pass as a psychological oddity, but Aristotle is referring to a change in the actual mirror.

be full of images and deficient in judgment, emotion enriching it in the first way and impoverishing it in the second. This is precisely Aristotle's conclusion. In sleep judgment is inhibited, the emotions have full sway, and the subliminal sensory activities, which are not noticed during the day because of the powerful trends of genuine perception, take on the character of reality. The images stream along in the moving blood, popping to the surface as the suppressing forces relax here and there, and pass for real so long as there is no awareness of being asleep. The images may continue to appear even after one has become aware of being asleep; but they are not then judged as real. The dream is an illusion. When it passes over into the perception of real external objects, as when one is affected by the light from a lamp and judges it as such and opens his eyes to find the lamp really there, it has ceased to be a dream.

Aristotle gives small credence to divinely inspired prophetic dreams. Dreams may be joined to the future, however, in three ways: as token, as cause, and as coincidence. The small sensory impressions which enter into dreams may betoken bodily changes, such as precede an illness; the images of the dream may be the starting-point, and hence the cause, of waking thoughts that issue in action; and finally, among the great variety of images which invade the dreaming mind, especially in the case of flighty persons, some are bound to coincide by chance with subsequent realities.

Plato is more favorably inclined toward dream divination than his dry-minded pupil. In the *Timaeus* it is anatomically provided for. But it is mainly in his discussion of the relation between dreaming and desire that he catches the attention of the modern reader. It is true that Aristotle speaks of dreams as conforming to the emotions of the dreamer; but he does it so incidentally, he is so much more concerned with the sensory and intellectual aspects, that the remark loses its effect. Not so with Plato. Just at the crisis of the discussion about the character of the tyrant, in the opening of the ninth book of the *Republic,* in a dramatically tense digression, Socrates is made to say that dreams reveal the violent and lawless desires which exist in every man. If this were all, it might be passed by as a mere apperçu. But the point is dwelt on very intently, very seriously, as offering a much needed ingress into the depths of human personality; and it is developed, though briefly, into a full theory of the dream. The omission of any reference to this theory in Freud's book is astonishing. Plato's theory is that when the rulership of reason is suspended during sleep, the other two elements of the soul, desire and the spirit of anger, reveal themselves in unmitigated viciousness. The soul, in this state, is ready to attempt incest, murder, and sacrilege. Even the most decent men express these criminal tendencies in their dreams. There are, however, morally superior dreams. These occur when the two lower elements of the soul have been lulled to rest by moderate satisfactions and reason has been stimulated to heightened activity. It is clear that Plato regards the dream as an expression of the dynamics of the total personality.

When we turn to the *De Rerum Natura* of the Epicurean atomist, Lucretius, we encounter a much less psychological theory.[8] Here the images are said to originate not in any operation of the mind but in the external world itself. The atomic combinations, whether shed from the surface of solid objects or generated spontaneously in mid-air, crowd continuously and fortuitously into

[8] *Cf. Of the nature of things,* a metrical translation by W. E. Leonard, London: Everyman's Library, 1921, pp. 162–174.

the mind of the sleeper. The illusion of movement in the dream is produced by a rapid succession of these static images: a cinematic effect. It is perfectly in keeping with Lucretius' theory that absurdities such as centaurs should crop up in the process; by chance-collisions various atomic patterns may fuse together into one strange form even before they appear to the dreamer. It is a much more serious problem how it happens that the dream-image so often conforms to the character of the dreamer and takes its place in a logical, or at least dramatically coherent, sequence. In a famous passage he adduces many examples of this fact. But he is not able to deal successfully, even in his own opinion, with the problem he raises. He considers the possibility that the will in some mysterious way commands the desired images, to come in a moment's space from wherever they may happen to be; and then turns dissatisfied from this answer to the one that every kind of image is present already wherever the dreamer may happen to be and enters consciousness when it is attended to, while those images which are not attended to drop into the unconscious background. Neither answer pleases him, and he dismisses the whole question as unnecessarily subtle. Indeed, he immediately warns the reader against looking for purpose anywhere in nature. All is chance. The human will, whose intrusion into the discussion seemed to ruffle his atomistic patience, is after all only a mechanical event following on the chance incursion of an atomic pattern into the temporary atomic organization which Lucretius calls the mind. Regardless of the philosophical adequacy of his position, Lucretius does an important service in presenting a vivid account of the theory that the elaboration of the dream takes place largely in the objective world outside the dreamer's consciousness—a view that harmonizes as well with the popular demon theory of his own times and later as it does with the atomic theory of Democritus.

II

Plato and Aristotle are the ancient precursors of Freud within the classic tradition. But neither occupies himself to any degree with the work of specific dream interpretation. That was an interest which demanded the practical turn of mind of a soothsayer or a physician. Such a man was Artemidorus of Daldis, a physician of the time of Antoninus Pius, whose *Oneirocritica* has served as a manual of dream interpretation and as a progenitor of popular dream books for centuries.[4] In the dedicatory epistle of the edition here used, published at Lyons in 1546, the Latin translator, Janus Cornarius, physician, apologizing for bringing out such a work in 'hoc turbulentissimo seculo,' notes that even in his own day Artemidorus was regarded with scant respect by the philosophers.[5] Cornarius feels impelled to support his case by citing Hippocrates, Galen, and the Bible on the importance of dreams.

Artemidorus distinguishes first of all between two classes of dreams: the *somnium* which forecasts the future, and the *insomnium* which refers to contemporary matters. The *insomnium* is affected by states of body, or of mind, or both. "The lover occupies himself with his sweetheart, the fearful man sees what he fears, the hungry man eats, the

[4] "Artemidorus' work was printed in Greek at Venice in 1518, in Latin at Basel in 1539 and 1544, at Lyons in 1546, in Italian at Venice in 1542, 1547 and 1548. The twentieth edition in English translation was reached in 1722 and the twenty-fourth in 1740." Harry B. Weiss, *Oneirocritica americana*, New York Public Library, New York, 1944, p. 6.

[5] Artemidori Daldiani philosophi excellentissimi de Somniorum Interpretatione Libri Quinque/A Iano Cornario Medico Physico Frankofordiensi Latina lingua conscripti/Lugduni apud Seb. Gryphium, 1546.

thirsty one drinks."[6] As for the *somnium*, it is sometimes a clear representation of coming events, sometimes symbolical. The general definition of *somnium* is "a movement or notion of the multiform soul, signifying good or evil in the future."[7] Every dream of important and serious men must be regarded as a *somnium*, because their souls are not disturbed by the hopes and fears and appetites which produce *insomnia*. Others, common people and dream interpreters alike, are subject to *insomnia*. The dreams of interpreters, however, tend to be obscured by their knowledge of symbolism. "The ordinary person sees in his sleep just what he desires or fears. But those who are skilled in interpretation discern their wishes through the veil of symbols. If indeed anyone who is uninformed in these matters has symbolical dreams they must be taken as *somnia*, not *insomnia*. For example: In the case of one who can understand symbols, either because he has chanced to read books on dream interpretation, or because he has dealt with dream interpreters, or because he has a talent in that direction, if it happens that he is in love with some woman, he will not see the object of his passion, but instead a horse, or a mirror, or a ship, or the sea, or a female animal, or a woman's garment, or some other thing which may represent a woman symbolically."[8] One may conclude that Artemidorus judges symbolical dreaming to be beyond the layman's normal range and to signify, therefore, in his case, the irruption of prophetic powers.

In general it is highly important to note whether the dream corresponds with the dreamer's ordinary ways. Whether it is prophetic of good things or bad, or whether it is prophetic at all, depends on the harmony or lack of it between the dream and the dreamer's daily life. There are six critical points to be considered: *natura, lex, consuetudo, tempus, ars, nomen*. That is to say, the interpreter wants to know whether the dream events are natural, lawful, and customary for the dreamer, what the circumstances were at the time, and what the dreamer's occupation and name are. Artemidorus' handling of some of these points, along with some very characteristic etymologizing, may be seen in the chapter entitled 'On Dreaming of Having One's Head Shaved,' which I translate here as an example of the method.

"To dream of having one's whole head shaved, except in the case of priests of the Egyptian gods, and those who study how to raise a laugh, and those whose custom it is to shave, for whom it is good, is in general a bad dream, because it signifies the same thing that nudity does, and indeed foretells sudden and dire misfortunes. To sailors it clearly portends shipwreck, and to the sick a most critical collapse, but not death. For the shipwrecked and men preserved from a serious illness are shaved; the dead, not at all. As to the aforementioned, it is a good dream because of their custom of shaving. To have one's hair cut by a barber, however, presages good to everyone equally. For the Greek word for barbering, κείρεως, is, I should say, from the word καρῆναι, which by the change of a single letter becomes χαρῆναι and takes on the meaning of joy or welcome; and surely nobody overtaken by any sort of misfortune or calamity has his hair cut, but only those who give much attention to their appearance and are free from sorrow and are not poor. But by a barber, be it noted, since if anyone dreams of cutting his own hair it means some sudden grief or calamity charged with great evils. For those who are preoccupied with such things have to do their own barbering. In addition, to have one's scalp scratched signifies cancellation of interest in the case of a debtor, but for others it means a loss through those by whom they have been scratched, if indeed they have dreamed of

[6] *Op. cit.*, Lib. I, Cap. I.
[7] *Ibid.*, Lib. I, Cap. II.
[8] *Ibid.*, Lib. IIII, Cap. I.

having their scalp scratched by others. For we say that a man has been 'scalped' if he has suffered a loss and been deceived by another person." [9]

It will be seen from these notes that Artemidorus appreciates the influence of desire in the formation of dreams, and that he is unwilling to apply his system of symbols without going into the circumstances of the dreamer's life. He particularly cautions the reader against the time-honored method of interpreting the dream by contraries. The sample dreams given in his Fifth Book have a very authentic appearance. His interpretations, however, waver curiously between farfetched prophesying and shrewd judgments of character. I translate two for the sake of example.

"An athlete dreamed that he became pregnant and gave birth to two black female children: he became blind, and his blinded pupils (*pupillae prolapsae*) became black." [10]

"A woman dreamed that her lover gave her the head of a pig. She conceived an aversion for her lover, and gave him up. For a pig is odious to Venus." [11]

III

A pagan author like Aristotle is free to deny the prophetic nature of dreams at his pleasure. Christian authors, faced by the numerous examples of prophetic dreams in the Bible, are in a different position. They are bound to admit that such dreams are possible. They are almost equally bound to develop a theory of dreaming which separates dream prophecy from rational prediction, since God, not the dreamer, is the author of prophetic dreams. On the other hand, the silliness and wickedness of many dreams will prevent the attribution of all without exception to the Divine Power. The exigencies of the Christian view require something different from Aristotelianism, or popular oneiromancy.

According to the authoritative Tertullian, writing about 203 A.D.,[12] sleep is a kind of death during which the always active soul is temporarily absent from its body. Dreaming is simply the activity of the soul deprived of its instrument.

"When, therefore, rest accrues to human bodies, it being their own especial comfort, the soul, disdaining a repose which is not natural to it, never rests; and since it receives no help from the limbs of the body, it uses its own. Imagine a gladiator without his instruments or arms, and a charioteer without his team, but still gesticulating the entire course and exertion of their respective employments: there is the fight, there is the struggle; but the effort is a vain one. Nevertheless the whole procedure seems to be gone through, although it evidently has not been effected. There is the act, but not the effect. This power we call *ecstasy*, in which the sensuous soul stands out of itself, in a way which even resembles madness. Thus in the very beginning sleep was inaugurated by ecstasy: 'And God sent an ecstasy upon Adam, and he slept.' [Gen. ii, 21.] The sleep came on his body to cause it to rest, but the ecstasy fell on his soul to remove rest: from that very circumstance it still happens ordinarily (and from the order results the nature of the case) that sleep is combined with ecstasy. In fact, with what real feeling, and anxiety, and suffering do we experience joy, and sorrow, and alarm in our dreams!" [13]

One might gather from the preceding that Tertullian thinks of the dream as something produced by the soul. But it appears that that is erroneous. The soul is a passive spectator, or at the most a reacting spectator; dream experience corresponds to realities over which the soul has no power. It is not the soul

[9] *Ibid.*, Lib. I, Cap. XXIII.
[10] *Ibid.*, Lib. V, Cap. XLIIII.
[11] *Ibid.*, Lib. V, Cap. LXXX.

[12] 'A treatise on the soul.' In *The ante-Nicene fathers* (Trans. by Peter Holmes), New York: Scribners, 1918, Vol. III.
[13] *Ibid.*, Ch. XLV, p. 223.

which produces the substance of the dream, but God or nature or (most commonly of all) the demons.[14] This theory has some kinship with the Democritean view exemplified in Lucretius.

Tertullian, however, is not consistent. He does not manage to believe without question that the soul is powerless to produce anything, even imaginations, and is thus always a pure mirror of consciousness held up to a world of realities independent of itself. But he tries hard, as if determined not to fall into the gulf of solipsism which opens beneath all those theories that assign a creative power of its own to the soul, whether that be the commonly assumed power of imagination, or the more exotic one by which projection of new realities into the external world is supposed to take place.[15]

A subtler and more systematic author of Christian persuasion is Gregory of Nyssa. Several passages in his treatise *On the Making of Man* (380 A.D.) [16] are devoted to our subject. In general, he treats the dream naturalistically, following Plato, Aristotle, and the physicians. But he distinguishes himself from them by his typically Christian insistence on the simplicity, unity, and independence of the soul, which for him is something different not only from the body but also from its own faculties of nutrition, sensation and reason. Also, he makes a place for prophetic dreams inspired by God; not, however, as Tertullian does, but by asserting that such dreams are miraculous occurrences, or rather that they are not dreams in the ordinary sense at all. "But if Daniel and Joseph and others like them were instructed by Divine power, without any confusion of perception, in the knowledge of things to come, this is nothing to the present statement; for no one would ascribe this to the power of dreams, since he will be constrained as a consequence to suppose that those Divine appearances also which took place in wakefulness were not a miraculous vision but a result of nature brought about spontaneously. As then, while all men are guided by their own minds, there are some few who are deemed worthy of evident Divine communication; so, while the imagination of sleep naturally occurs in a like and equivalent manner for all, some, not all, share by means of their dreams in some more Divine manifestation." [17]

Except for these occasional miraculous communications from the Divine, dreams are considered by Gregory as belonging to the order of nature and subject to psychological explanation. Sleep, he says, is a condition in which the sensory faculty of the soul is at rest; and along with that, the intellect also; for the intellect, or mind, or rational faculty, is conjoined with the senses. The absurdity of dreams is evidence, indeed, that reason has nothing to do with them. Dreaming, then, must be assigned to the third main faculty of the soul, the nutritive, which alone remains active during sleep. It is not a function of the nutritive faculty, of course, to produce images: these have fallen into it by chance, as it were, from the waking operations of the higher faculties. The soul is however not completely transformed by sleep. There is simply a reversal of the waking hierarchy of faculties, so that the two higher faculties are dominated by the nutritive.

So much for Gregory's theory of the dream in general. When he comes to deal with dream content, he has a number of penetrating observations to make, some of them original with himself. The

[14] *Ibid.*, Ch. XLVII.

[15] *Cf.* Richard Wilhelm and C. G. Jung, *Das Geheimnis der goldenen Blüte*, Munich: Dornverlag Grete Ullman, 1929.

[16] In *A select library of Nicene and post-Nicene fathers of the Christian church.* New York: Scribners, 1917, Second series, Vol. V.

[17] *Op. cit.*, Ch. XIII, 12, pp. 400f.

dream content, he says, has reference either to memories of daily occupations, or to the state of the body at the time. He repeats what many others have said, that "the thirsty man seems to be among springs, the man who is in need of food to be at a feast, and the young man in the heat of youthful vigor is beset by fancies corresponding to his passion." [18] But he adds to all this a personal observation that catches sight of the connection, in a concrete instance, between bodily condition and dream imagery. He writes: "I also knew another cause of the fancies of sleep, when attending one of my relations attacked by frenzy; who being annoyed by food being given him in too great quantity for his strength, kept crying out and finding fault with those who were about him for filling intestines with dung and putting them upon him: and when his body was rapidly tending to perspire he blamed those who were with him for having water ready to wet him with as he lay: and he did not cease calling out till the result showed the meaning of these complaints: for all at once a copious sweat broke out over his body, and a relaxation of the bowels explained the weight in the intestines. The same condition then which, while his sober judgment was dulled by disease, his nature underwent, being sympathetically affected by the condition of the body—not being without perception of what was amiss, but being unable clearly to express its pain, by reason of the distraction resulting from the disease—this, probably, if the intelligent principle of the soul were lulled to rest, not from infirmity but by natural sleep, might appear as a dream to one similarly situated, the breaking out of perspiration being expressed by water, and the pain occasioned by the food, by the weight of the intestines." [19] Gregory relates this striking observation to the general view of the physicians who say that their patients reveal the nature of their complaints in their dreams, making possible the differentiation of stomach ailments from cerebral, etc. Besides, dreams reveal the dreamer's general character, distinguishing between the brave and the cowardly, the lascivious and the continent, the liberal and the avaricious.

Finally, a word must be said about Gregory's view of the passions, since they, in his opinion, contribute to the formation of dreams. The fountainhead of all the human passions, he says, is sexual reproduction. He seems to mean by this two things: first, that sexual desire is the fundamental passion; and second, that through this manner of reproduction man is held down to the level of brute life and shares in its nature. "For those qualities with which brute life was armed for self-preservation, when transferred to human life, become passions; for the carnivorous animals are preserved by their anger, and those which breed largely by their love of pleasure." [20] But the animal tendencies are multiplied and intensified in man by the action of reason itself, which is corrupt through sin. "Thus our love of pleasure took its beginning from our being made like the irrational creation, and was increased by the transgressions of men, becoming the parent of so many varieties of sins arising from pleasure as we cannot find among the irrational animals. Thus the rising of anger in us is indeed akin to the impulse of the brutes; but it grows by the alliance of thought: for thence come malignity, envy, deceit, conspiracy, hypocrisy." [21] The reader will sense in these words a foreshadowing of more than one modern author.

The last ancient writer on dreams to be considered here is Synesius of Cyrene, once a pupil of the famous Neo-

[18] *Ibid.*, Ch. XIII, 14, p. 402.
[19] *Ibid.*, Ch. XIII, 15, p. 402.
[20] *Ibid.*, Ch. XVIII, 2, p. 408.
[21] *Ibid.*, Ch. XVIII, 4, p. 408.

Platonist Hypatia, and later the somewhat unorthodox Bishop of Ptolemais, who wrote his *De Insomniis Liber* in the latter part of the fourth century.[22] The work was edited and translated from the original Greek into Latin by Dionysius Petavius early in the seventeenth century. Before that, in the fourteenth century, it had received a lengthy commentary by the Byzantine historian Nicephorus Gregoras.[23] Apparently it may be considered a work of some influence for a period of several hundred years.

Synesius believes in divination generally: the world is everywhere connected, and happenings in one part may set up sympathetic vibrations in another part. Dreams are especially to be recommended because they are available to all men, and because, unlike other means of divination, they are private and cannot be laid under a legal ban. They are obscure, but no more so than oracles.

The full understanding of dreams requires an excursion into psychology. The *mind* contains the forms of existing things; the *soul*, the forms of things which are coming into being. The contents of the soul are known to us through the mirror of *phantasy*, and only thus. Phantasy is soul's first body; it is the common center of all the senses, and from the center rules over them as its instruments. It is a far purer sense than any which rely on bodily organs, as, for instance, the eye, whose accuracy is hindered in several ways. Phantasy is not limited to man; it is found widespread in nature, in irrational animals and in demons.[24] In man, though it may work alone, it is more usually found operating in conjunction with mind. To transcend phantasy completely is a difficult and blessed achievement. In ordinary life one lives under its dominion and through it suffers punishment. Some are rising, some descending in the scale of intelligent knowledge, those being in the worst state who do not recognize the evil of their present lives and do not desire to escape from it. The will to escape from evil is the most essential step in the whole process of purification. The soul's position in the scale is revealed by the nature of its phantasy. It is the punishment of evil men that their phantasy is evil; the reward of the good that theirs is good.

As already noted, Synesius regards phantasy as the means by which the soul represents to consciousness its knowledge of things to come. The role of phantasy is further clarified in a passage where he discusses his whole theory of knowledge. I translate: "For this is the way the matter stands: whatever things there are in nature, of all that is or has been or shall be, according to the plan of existence, images flow off and leap forward from their substance. For

[22] Synesii De insomniis liber. In *Patrologiae cursus completus*. Migne. Series graeca. Tom. 66. 1864. Cols. 1281–1320.

[23] Nicephori Gregorae Explicatio in librum Synesii de insomniis. In *Patrologiae cursus completus*. Migne. Series graeca. Tom. 149. 1865. Cols. 520–642.

[24] I translate here a part of Gregoras' note. "All the classes of demons, he says, which inhabit this world, have been allotted that sort of life which has the power of sensation. For, in a word, those classes of demons are simply imaginations. . . . Images, he says, though they are destitute of body, nevertheless have a kind of density, some more, some less. Here Synesius appears to be following the opinion of Democritus and Sextus and other Greeks. For they assert that there are certain demons round about the earth and the air, having indeed density and coherence, invisible to human eyes (if in normal health), which they call spectres, because, I think, they retain and conserve a kind of form and empty appearance and likeness of angelic and immaterial nature, from which they have fallen, and difficult to destroy, though not perpetually immortal, many of which are harmful and malignant, though some are beneficent. These, they say, draw near to men, and attach themselves to them, and especially to those who are asleep, and prophesy to them many things which are to come to pass." *Op. cit.*, Col. 567.

just as each of those things perceived by sense is form combined with matter, so also is the flux of matter comprehended by us as patterned; reason itself convinces us that images are by nature derivative, since in every case those things which have a beginning must degenerate from the dignity of that which is. To all these flowing images the spirit of phantasy is the most perfect mirror. For when, fallen from their previous state and blindly wandering, they, because of the instability of their nature and because they are known by nothing, enter into living spirits, which are in general images but with a fixed and settled position in nature, they cling to them and come to rest in them as in a secure dwelling-place. From things past, clear images are sent out, until little by little in the course of time they fade away; but from things present, inasmuch as they still abide, the images are clearer and more vital; while, finally, those coming from things future are relatively uncertain and indistinct. For there are antecedent unfoldings of things not yet present, germs of the imperfect nature leaping out as it were, and hints of the hidden seeds breaking forth." [25]

For Synesius, then, there exist theoretical reasons for developing a science of divination. But though he wishes to establish such a science, a science of images which would enable him to predict from his dreams what was about to take place in reality, as a pilot may recognize the port he is nearing by the rocks which stand at its entrance, he pours scorn on the numberless books in his time which set forth general laws of relationship between dream image and future event, because, he argues, every person's phantasy has its own personal nature, and a variable one at that. He says of those who have attempted to develop a system of standard relationships that he is convinced that they have never reflected philosophically on the nature of spirit at all. Nevertheless, it may still be possible to discover the laws governing one's own phantasy, insofar as it is constant, by keeping a careful record of one's dreams and the events which follow. This he himself has undertaken to do. But he adds that it is tedious work, trying to express accurately in words the wild illogicality of dreams.

[25] *Op. cit.*, Col. 1310 f.

IV

The hope that originally inspired this essay, that a search through ancient and medieval writings, particularly medieval, might turn up some valuable new insights into the nature of the dream, has been largely disappointed. A fairly patient consultation of the volumes in the Migne collection yielded little of a speculative nature beyond the works by Synesius and Nicephoras Gregoras. There is, however, this one interesting general conclusion to be drawn, historically: that modern dream theory, as represented particularly in Freud, is in most respects closer to Aristotle and Plato than it is to the later writers examined here. Especially remote from modern speculations is the atomism of Lucretius (and the collateral demonism of Tertullian). One might find occasion for surprise in this fact, in view of the increasing dominance of science by the atomism of modern physics. Just about equally remote, however, is the standpoint of a writer like Synesius with his Neo-Platonic emphasis on the creativity of the soul.

AGE AS A FACTOR IN THE RECALL OF INTERRUPTED TASKS

BY R. NEVITT SANFORD

University of California

An experiment that was worth doing is usually worth repeating. This truism was brought home to the writer by reading in the PSYCHOLOGICAL REVIEW Prentice's recent critical paper on the influence of task-interruption upon later behavior. Prentice (2) mentions numerous confusions and contradictions that have arisen since Zeigarnik's original contribution, and suggests that theoretical treatments have failed to keep pace with the experimental evidence. While granting that in certain areas this has been the case, one may suggest that in other areas theory has run ahead of the facts. Thus, certain crucial features of contemporary theory concerning interruption, as summarized by Prentice, seem still to rest entirely upon a pioneer experiment performed by Rosenzweig in 1933, while a repetition of this experiment has yielded different results having different theoretical implications. Overgeneralization from Rosenzweig's results seems to have pyramided with the passage of time, and effort spent in trying to reconcile them with expanding theory might have been saved by a check upon their general validity.

The findings in question refer to the relation between age (both C.A. and M.A.) and the tendency differentially to recall completed and interrupted tasks. Rosenzweig's subjects were 40 crippled children ranging in age from 5 years and 6 months to 14 years and 8 months. They worked with a series of jig-saw puzzles on half of which they were failed (interruption) and on half of which they were allowed to succeed (completion). Rosenzweig reports that "the group who recalled successes better than failures was differentiated from the group who recalled failures better than successes by a more advanced mental age . . ." (3, p. 480). The writer, repeating Rosenzweig's procedure with 49 normal children, found that "the tendency to remember failures better than successes increases with chronological age during the years 7–15 and with mental age in the range of 10 years to 21 years" (5).[1] Rosenzweig, evaluating his above result in relation to other findings, interpreted them "as indicating that repression is a mechanism of defense resorted to relatively late in the development of the child" (3, p. 481). The present writer, on the other hand, concluded that the tendency to recall failures better than successes is an expression of the age-linked factor of ego strength—and that forgetting failures is preëminently a childish mode of defense.

Before attempting to rationalize these findings and conclusions it is necessary to examine in detail the results of these two studies themselves. We must consider not only C.A. and M.A. but also the age-linked factors of pride, self-criticism, perseveration, and the tendency to resume the interrupted tasks.

The writer's procedure followed that of Rosenzweig and Mason (4) very closely. He used the same puzzles, the same order of successes and failures, and the same procedure for obtaining recall. However, there were certain differences

[1] These results were published only in summary form and in a manner highly incidental to other material, so Prentice can certainly be excused for overlooking them. The complete data are presented now for the first time. Their relevance to an important and complex problem seems more apparent than before.

in the experimental conditions which may be of some significance. Whereas Rosenzweig's subjects were crippled children living in an institution, the writer's subjects were normal, healthy children in a private school—a better group, it would seem, upon which to base generalizations about the effects of task-interruption in children. In order to insure high motivation Rosenzweig offered a prize to the subjects obtaining the best score, while the present writer merely instructed his subjects, "see how well you can do." These private school children lived in an atmosphere where intellectual achievement was highly valued; they were highly ambitious as a group and they took pride in their intellectual accomplishments. Suspecting that failure on the puzzles would go hard with them, the writer allowed one more of the successes than failures, whereas Rosenzweig had allowed the same number of successes and failures.

The puzzles were presented to 49 children in 1936 (Puzzles I); a different set of puzzles was presented to 26 of these same subjects 2½ years later (Puzzles II).

1. Chronological age. In Fig. 1, the per cents of subjects who recalled failures better than successes, in Puzzles I and in Puzzles II, have been plotted against C.A. Rosenzweig and Mason's graph of this same function is also included. The results from both Puzzles I and Puzzles II seem clearly to show that with increasing C.A. subjects recall more failures. Rosenzweig and Mason found in their results no clear indication of such a tendency. From an inspection of Fig. 1, however, we are led to question the reliability of the final point on their curve (13–14) and to say that with this exception their findings concerning C.A. resemble the present ones.

2. Mental age. The relation of M.A. to preference for failures in recalling puzzles (Puzzles I, Puzzles II, Rosenzweig and Mason's data) is shown in Fig. 2. Both Puzzles I and Puzzles II seem to show that preference for failures

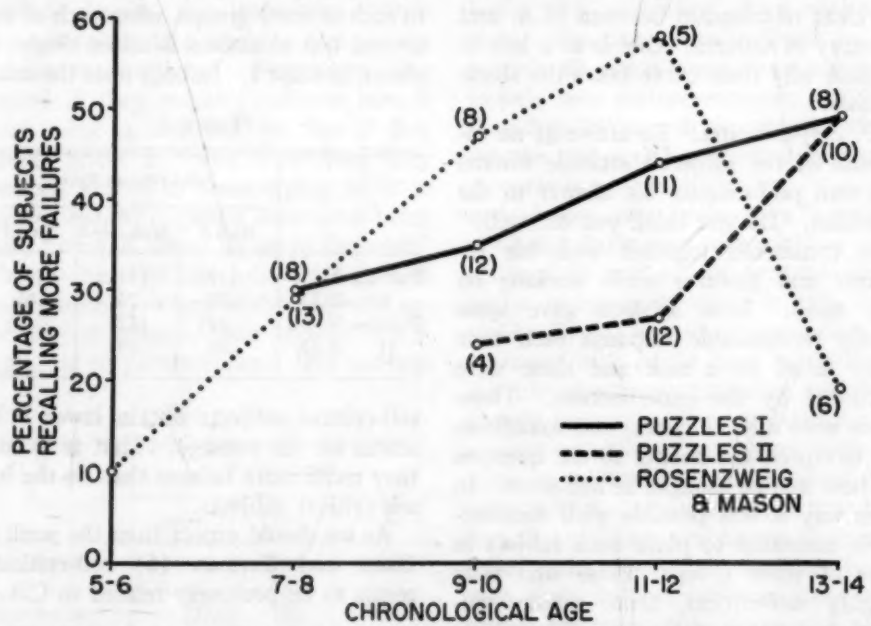


FIG. 1. Relation of chronological age to preference for failure in recall. Figures in parentheses indicate the number of subjects in each mental age group.

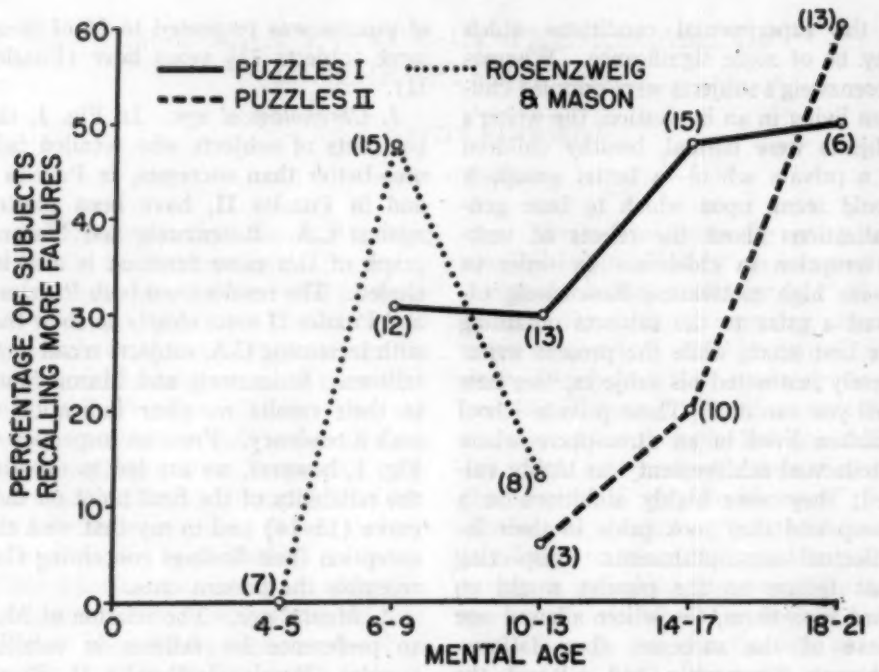


FIG. 2. Relation of mental age to preference for failure in recall.

in recall increases with M.A. The results from Rosenzweig and Mason show no clear relationship between M.A. and memory of failures. One is at a loss to explain why their curve takes the shape it does.

3. Self-criticism. To arrive at an estimate of the subject's attitude toward his own performance his answer to the question, "Do you think you did well?" was considered together with his remarks and gestures while working on the tasks. Most subjects gave some easily recognizable response each time they failed on a task and these were recorded by the experimenter. These data were used to check, and sometimes to interpret, the answer to the question of how well he thought he had done. In this way it was possible with considerable assurance to place each subject in one of three groups: those who were highly self-critical, those whose self-criticism was moderate or indeterminate and those who showed little or no self-criticism. The average recall difference (R.D.) score [2] of the subjects belonging to each of these groups, when each of the several test situations is taken singly, is shown in table I. In both tests the more

TABLE I

		Self-Criticism Rating		
		High 3	Moderate 3	Low 3
Puzzles I	N	(17)	(11)	(3)
	R.D.	−.03	.24	.46
Puzzles II	N	(7)	(12)	(7)
	R.D.	.06	.19	.24

self-critical subjects obtain lower R.D. scores on the average. That is to say, they recall more failures than do the less self-critical subjects.

As we should expect from the work of Binet and Terman (6) self-criticism seems to be positively related to C.A.—

[2] $\text{R.D.} = \frac{\text{successes recalled}}{\text{successes given}} - \frac{\text{failures recalled}}{\text{failures given}}$

$r = .43$ (Puzzles I) and .25 (Puzzles II)—and to M.A.—$r = .33$ (Puzzles I) and .36 (Puzzles II). These results seem to offer some explanation of the findings concerning memory of failures in relation to C.A. and M.A. Self-criticism, which is associated with the better recall of failures, increases with C.A. and M.A. and so does the tendency to recall more failures.

Rosenzweig and Mason found a positive relationship between C.A. and self-critical answers to the question, "Do you think you did well?" but there seemed to be no relation between the latter factor and memory of failures. This latter finding does not necessarily contradict the present ones for, as a glance at Fig. 2 will show, we are here dealing with higher mental ages. Probably, in view of its correlation with mental age, we are also dealing with higher levels of self-criticism.

4. Perseveration on unfinished puzzles. Some of our subjects insisted on working on the puzzles after time was called, despite the instruction "when time is called, stop working and put the pieces in the box." Sometimes they asked if they might "just see how it was going to turn out" or "see if they could finish it," and sometimes they seemed unable to resist trying to fit in one or two more pieces even after being told "we had better go on to the next." On the basis of the behavioral data it was possible to group the subjects according to whether they showed much, little, or no perseveration. The average R.D. scores of subjects belonging to each of these groups are given in table II. It appears that in both the Puzzles I and the Puzzles II tests there is an association between the tendency to perseverate on unfinished puzzles and the tendency to remember failures better than successes.

TABLE II

		Perseveration		
		Much	Little	None
Puzzles I	N	(10)	(10)	(12)
	R.D.	−.10	.25	.14
Puzzles II	N	(7)	(8)	(11)
	R.D.	.14	.19	.25

Perseveration on unfinished puzzles seems to be positively associated with C.A.—$r = .15$ (Puzzles I) and .36 (Puzzles II); with M.A.—$r = .25$ (Puzzles I) and .39 (Puzzles II).

We may now consider the implications of the results just reported for the general theory of task interruption. Our results bear most directly upon those parts of the theory which have to do with maturity and the type of personality organization that goes with it. Prentice, in his summary of Rosenzweig's work, says that "the mature and the proud tend to 'repress' the memory of their failures, as a result, presumably, of a strong need for success or for social status." Just on a common sense basis, this statement does not seem right. Reacting to one's failures by forgetting all about them does not seem to be a very mature response, any more than does letting a drive for social status get one into such a state that he has no recourse but to distort reality.

And why, one might ask, if 'repression' of failures goes with a higher mental age do not adults overwhelmingly prefer successes in recall in this situation? Repetitions of Rosenzweig's experiment with adults have never shown a truly significant majority preferring success in recall. The argument for the greater maturity of repressers rests not alone upon Rosenzweig's statement concerning M.A. but on a further report by him that children who choose to *resume* failed tasks *recall* their successes better than their failures, whereas those who repeat the previously successful tasks remember mostly failures. Resumption

of interrupted (failed) tasks does clearly increase with age (Rosenzweig)—and if those who resume are those who repress, then repression must also increase with age. Now the present results state pretty flatly that repression *does not* increase with age (either C.A. or M.A.); and one may use them to call into question Rosenzweig's assumption of an inverse relationship between resumption and recall. According to Lewinian theory, which has been rather well supported, a tendency to resume an interrupted task and a tendency to think about it or recall it are expressions of the same tension system. This seems reasonable. For instance, why should a person, unless he be masochistic, dwell upon his failures unless he intended to do something about them when the chance arose? This would seem to be what was going on in a sizeable group (24 per cent of the 35 subjects) in an experiment by Cartwright. These subjects, asked to rate a series of tasks in order of preference, raised their preference-rating after failure. "Interviews with the subjects seemed to indicate clearly that many subjects who raised their preference-rating received the experimenter's evaluation of their performance not as failure, but rather as an obstacle in the way to the goal. . . . Other subjects seemed to want to 'erase' the failure by subsequent success in the same activity" (1, pp. 11–12). This is almost exactly the way the present writer would describe his subjects who preferred failures in recall. They were critical of their performance, they blamed themselves for their failures or at least they took a certain responsibility for them. They perseverated markedly on unfinished puzzles and gave every indication that if they had a chance to finish them later they would certainly make the most of such an opportunity. Prentice (2, 331) writes further that "In terms of Rosenzweig's formulation, Cartwright's subjects who preferred tasks on which they had previously failed would be those with strong 'ego defenses.' *They are presumably the mature and prideful subjects*" (italics mine). In other words, in terms of Rosenzweig's formulation, they are the subjects who repress. This mixes things up badly. The writer would argue from his own results, as well as from the Lewinian experiments and from clinical observation, that these subjects of Cartwright's are precisely the subjects who would *not* repress, that they are indeed relatively mature and have strong 'ego defenses' which do not include repression as a major one, and that the pride which they have is of a different sort than that which teachers ascribed to Rosenzweig's subjects who preferred success in recall.[8]

In summary of his interpretations of Rosenzweig's and Zeigarnik's work Prentice writes, "any threats to the subject's view of himself will produce a tendency to forget those threats but also to overcome them. The mature subject or the subject who has sufficient pride will be more sensitive to such threats, and 'important' tasks will be more threatening when failed than will incidental ones" (2, 331). What needs to be stressed here is that forgetting the threats to self-esteem is one device for self-defense and attempting to overcome one's failures is a totally different one. Thus, although both mechanisms might well be employed by the same individual, most individuals probably tend to employ the one or the other most of the time. With maturity, in our subjects at least, there is an increasing use of 'overcoming' and a decreasing use of

[8] In Rosenzweig's instructions to the teachers who rated his subjects, pride is defined as 'eagerness to do well.' It would seem that this trait, which might spring either from an achievement drive or from a need for recognition, is not the same thing as self-respect. And it is quite different from the frame of mind which leaves a person 'too proud to retreat.'

'forgetting.' Thus it is true that subjects who attempt to erase their failures have the strong ego defenses, if by strong ego defenses is meant the relatively adequate defenses which we expect to find in a strong ego. Repression is the device characteristically resorted to by the weak ego, which attempts to solve its problems simply by not facing them. *A strong need to defend the ego* and *strong ego defenses* are two entirely different concepts. The former is a matter of motivation, involving narcissism, insecurity, etc., while the latter is a matter of technique or 'mechanism' for dealing with one's problems; the 'strength' of the former is a matter of motivational intensity, while the 'strength' of the latter lies in its adequacy or effectiveness. It is the former—a strong need to defend the self (or what is nowadays called ego involvement)—which in Rosenzweig's subjects seemed to be associated with more advanced mental age and with the higher teacher's ratings on pride. Probably a certain minimum of intellectual development and a certain degree of 'eagerness to do well' (pride) is necessary for a subject to see in Rosenzweig's task situation a threat to his self-esteem. But it seems wrong to call this condition 'maturity'; and it is even worse to assume that whenever such a state of affairs exists we necessarily get repression.

One can account in part for the discrepancy between Rosenzweig's results and ours on the basis that rather different levels of mental age or maturity were represented in the two groups of subjects. As shown in Fig. 2, Rosenzweig's subjects have a considerably lower average mental age than the present ones—in fact there is only a small overlapping between these two groups. It is apparent too, from a consideration of Figs. 1 and 2, that Rosenzweig's subjects are, as a group, relatively retarded in intellectual development. There is a question of what kind of personality organization one might expect most commonly to find among crippled children living in such a 'home,' but it would certainly not be surprising if their ego development were considerably less advanced than that of normal children of the same age. It may have been in anticipation of some difficulty in getting the subjects ego-involved that Rosenzweig set up a competition by offering a reward for the best performance. In any case, it seems that some of his subjects were ego-involved and others not, and that it was among the former that he found the repressers and the 'resumers.' It is the writer's opinion that *all* of his own subjects were ego-involved, and that variations in their eagerness to do well were not so much associated with C.A. as they were conditioned by numerous personality and situational factors. Being ego-involved, they reacted to failure according to their personality organization, the less mature ones trying to put a good face upon things by forgetting their failures and remembering their successes, the more mature ones trying to stick to reality (which would decree that they recall successes and failures evenly), accepting responsibility for their failures and thinking of how to erase them, even remembering failures as an expression of their tendency to keep in consciousness their conflicting attitudes toward themselves.

This does not explain away Rosenzweig's inverse relation between resumption and recall, though it would seem to cast some doubt upon the general validity of his findings. Now let us examine his results themselves.

In the experiment mentioned above, Rosenzweig found that 20 of 37 subjects chose to resume their successes and 17 chose to repeat their failures. Of the 20 choosing to repeat successes, 7 recalled more failures, 9 recalled evenly, and 4 recalled more successes. Of the 17 choosing to repeat failures, 5 recalled more failures, 1 recalled evenly, and 11

recalled more successes. These differences are clear enough, but N is so small that one would ordinarily reckon seriously with the possibility that chance factors might have had a large effect. Thus, we note that only 6 of the 20 subjects choosing to repeat successes were given more than 8 puzzles, too few, it would seem, to reveal a general trend in forgetting. We note further that of the 9 subjects in this group who recalled evenly, *6 were given only 6 puzzles.* We can presume that here most of the subjects recalled evenly because they recalled all the puzzles that were given. Thus the results for the group who chose to repeat their successes seem at best rather unsubstantial.

Results are more clear cut in the case of the 17 subjects who chose to repeat their failures. Here the age factor in resumption seems undeniable. The highest mental age found in the group choosing to repeat successes is 8 years, 5 months, while all the subjects with mental ages above 8 years, 5 months, chose to repeat their failures. Following Rosenzweig's lead we may consider that the 20 very immature children who chose to repeat their successes were not really involved in the task situation, while the 17 more mature subjects who chose to repeat their failures were so involved. In the latter group a clear majority recalled successes better than failures, presumably because their greater eagerness to do well led to greater involvement in the situation, and hence to a greater need to defend their self-respect. Now if this were all the data we had at our disposal we might well conclude with Rosenzweig that with increasing mental age there is increasing preference for successes in recall. The data presented in this paper, however, show that when we have to deal with higher levels of mental age—levels which begin near the place where those of Rosenzweig's subjects leave off, there is an increasing tendency to remember more failures. Surely we cannot attribute this to decreasing involvement in the task, because lack of involvement has already been ascribed to immaturity. We must, it seems, fall back upon the explanation that the more mature subjects were goaded by their pride into remembering their failures precisely because they hoped somehow to erase them or make up for them. Since *all* the subjects in the present experiment had M.A.s well above 8 years and 5 months it can be inferred, from Rosenzweig's data, that they *all* would have chosen to resume their failures. Thus the use with these subjects of Rosenzweig's procedure for obtaining repetition choice would hardly have suggested an inverse relation between resumption and recall! To check on the question as to whether in mature subjects the tendency to resume and the tendency to recall failed tasks are manifestations of the same tension system, it might be well to use a technique similar to Cartwright's—*i.e.*, estimating the *attractiveness* of completed and failed tasks—in conjunction with the present technique for obtaining recall. Until this, or a similar experiment, has been done, it cannot be said that the facts have even caught up with the theories, let alone surpassed them.

REFERENCES

1. CARTWRIGHT, D. The effect of interruption, completion, and failure upon the attractiveness of activities. *J. exp. Psychol.*, 1942, 31, 1–16.
2. PRENTICE, W. C. H. The interruption of tasks. PSYCHOL. REV., 1944, 51, 329–340.
3. ROSENZWEIG, S. The experimental study of repression. In *Explorations in personality.* Murray, H. A., *et al.* New York: Oxford Univ., 1938, 472–490.
4. ——, & MASON, G. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, 24, 247–265.
5. SANFORD, R. N., *et al.* Physique, personality and scholarship. *Monogr. Soc. Res. Ch. Devel.*, Vol. VIII, No. 1, 1943.
6. TERMAN, L. M., *et al. Genetic studies of genius.* Stanford Univ., 1925.

THE MEANING OF EXPLANATION

BY DAVID L. MILLER

University of Texas

I. Introduction and Background

Can we make a significant distinction between description and explanation? This question is foremost with modern scientists and philosophers who believe that a teleological approach to an understanding of the world is futile. In what follows we will make an attempt to say yes to this question and to show what explanation rightfully means in modern science.

Obviously the ancient and mediaeval belief that an understanding of natural processes requires a foreknowledge of that for the sake of which they take place passed away with the advent of the early modern scientists such as Galileo, Pascal, Descartes, Hobbes, etc. The question *why* do things happen as they do was superseded by the question *how* do they happen. Final cause (teleology) gave way to efficient cause (mechanics). Does this mean that explanation and understanding were supplanted by description? Must we, to be logically consistent, conclude that description is the only valid function of science and that our quest for explanation is a vestige whose fulfillment science forever precludes? I think not. But rather, although scientific explanation can never consist in finding final causes or that for the sake of which things take place, explanation gives an intelligent, rational account of what would otherwise be logically and rationally disconnected sense experiences. Certainly the purpose or purposes for phenomena, if meaningful at all, are not to be sought after in science. Admittedly teleological notions are gratuitous in the sense that, through over-generalization, they have been furnished by men's minds and superimposed upon the external world at the expense of a more significant understanding of natural processes. Purpose in this sense is a metaphysical concept having no possible confirmation in the science laboratory. Such a concept not only transcends experience, but, as August Comte and apparently all positivists since his day including many contemporary logical positivists are prone to conclude, any statement which goes beyond the facts (*i.e.*, any statement other than a *descriptive* one) is to be condemned as metaphysical and as a hindrance to science (1). A corollary to this conclusion is that the rightful function of science is to describe only, and that 'explanation' is significant of the psychological reactions of the 'explainer,' but contributes nothing to the body of scientific knowledge. Let us show that this conclusion and its corollary are false, and that, contrariwise, description without explanation is blind and sterile; admitting, of course, that explanation apart from description is incomplete and insignificant.

Apparently Karl Pearson assumes that all types of explanation are teleological in character and, consequently, if one rejects teleology as a scientific principle then science can do nothing other than describe. He says: "The scientific law is a description, not a prescription" (7, p. 104). Also, "Science for the past is a description, for the future a belief; it is not, and never has been, an explanation, if by this word is meant that science shows the *necessity* of any sequence of perceptions" (7, p. 136). And again, "When, therefore, we say that we have reached a 'mechanical explanation' of any phenomenon, we

only mean that we have described in the concise language of mechanics a certain routine of perceptions. We are neither able to explain why sense-impressions have a definite sequence, nor to assert that there is really an element of necessity in the phenomenon" (7, p. 140).

Such men as Pearson, Mach, and Poincaré object to the metaphysical contentions of science; they reject the belief that sensations (which to them constitute the sole object of scientific knowledge) have causes (things in themselves) which cannot be sensed. They grant, of course, that if there were things in themselves serving as a condition for sensations, and if we could know them, then we would be able to predict the order of sensations (or perceptions) through an analysis of their causes (things in themselves) and we would have both necessity and explanation. For the most part the rebellion against explanation has been of this general sort; it is a rebellion against the belief that there is a metaphysical (beyond sensation) condition for change, whether that condition be a purpose in the mind of God or a more natural one such as the "inner nature" of matter. In general, also, the attack on metaphysical beliefs has been made by showing that predictions made by way of an analysis of either purpose or matter are at best highly probable and never certain, as they should be if the future is an unfolding of an inherent necessity, be it in nature or in God's will.

Let us grant that Pearson and others are correct in rejecting teleology or any other 'inner necessity' as a scientific basis of explanation. But we cannot grant that "scientific law is a description." Rather, we shall contend, a scientific law can be and in most cases is an explanation, not a description. In order to avoid an argument over words we must show clearly the distinction, and finally the relationship, between description and explanation as they apply to modern scientific procedure.

Every explanation must have essentially two differentiating factors which description lacks.

(1) Every explanation must make a prediction.

(2) Every explanation must make the phenomenon explained intelligible by referring to more basic or more fundamental experiences which the person to whom the phenomenon is being explained will grant as psychologically (or intuitively [5, p. 1]) satisfactory. Let us discuss the first factor before discussing the second, and independent of the second.

II. DESCRIPTION *versus* EXPLANATION

August Comte wrote:

"Our business is,—seeing how vain is any research into what are called *Causes*, whether first or final,—to pursue an accurate discovery of these Laws, with a view to reducing them to the smallest possible number. By speculating upon causes, we could solve no difficulty about origin and purpose. Our real business is to analyse accurately the circumstances of phenomena, and to connect them by the natural relations of succession and resemblance. The best illustration of this is in the case of the doctrine of Gravitation. We say that the general phenomena of the universe are *explained* by it, because it connects under one head the whole immense variety of astronomical facts; exhibiting the constant tendency of atoms towards each other in direct proportion to their masses, and in inverse proportion to the squares of their distances; whilst the general fact itself is a mere extension of one which is perfectly familiar to us, and which we therefore say that we know;—the weight of bodies on the surface of the earth. As to what weight and attraction are, we have nothing to do with that, for it is not a matter of knowledge at all. Theologians and metaphysicians may imagine and refine about such questions; but positive philosophy rejects them. When any attempt has been

made to explain them, it has ended only in saying that attraction is universal weight, and that weight is terrestial attraction: that is, that the two orders of phenomena are identical; which is the point from which the question set out" (2, pp. 701–702).

From Comte's antimetaphysical doctrine springs the general conclusion that the sole function of science is to describe, not explain. But suppose we do nothing more than 'state the facts' in their spatio-temporal order. Suppose we do not go beyond the facts but let them 'speak for themselves,' lest we introduce extraneous, unwarranted metaphysical elements. What would this amount to, if it is possible at all? Is it possible for men to detect events in their native form, uninterpreted? C. I. Lewis, in *Mind and the World Order*, has shown us that it is not only impossible to do so, but that it is undesirable and even insignificant. But if we could submit an 'unbiased,' 'objective' record of the facts, that record would not have either a spatial or a temporal reference system involved. We could not, in submitting such a record, state any sort of relationship between the facts, such as before, after, simultaneous with, above, below, to the right or left of, etc. Our quest for sheer description would break down unless we resorted to the insignificant task of portraying images in a disorderly, disconnected, dream-like manner.

But suppose we permit one to describe the facts with reference to other facts; *i.e.*, in their spatio-temporal setting. This alone would never give us laws. We would have nothing more than statements about particular experienced events. We could never say, for example, that if objects are sliding down inclined planes, whose friction is negligible, the acceleration of descent is proportional to the sine of the angle (acceleration $= k \sin x$). Rather we would confine ourselves to particular instances. Pure description does not give us a law of nature. A law must always go beyond the given and, of course, description is limited to the experienced object, or, possibly, in *sheer* description, to the given. Karl Pearson writes:

"The object served by the discovery of such scientific laws is the economy of thought; the suitable association of conceptions drawn from stored and correlated sense-impressions, permits the fitting exertion to follow with the minimum of thought upon receipt of an immediate sense-impression" (7, p. 93).

But Pearson will never grant that laws, if valid, transcend a description of sense-experience. For him it is a summary of past experiences. ". . . it is a *résumé* or *brief expression* of the relationships and sequences of certain groups of perceptions and conceptions, and exists only when formulated by man" (7, p. 98).

Clearly if a law were confined to sheer description it could never make a prediction. In fact it could never state the relationship between events or facts, if those relationships were simply spatio-temporal ones. Past experiences and past events confined unto themselves can never legislate over the future. To assume that a description of causally disconnected 'facts' will enable us to foresee future events is gratuitous. Certainly extreme positivism, with its quest for *warranted* knowledge based on experienced facts, breaks down as a method of accounting for laws which in fact enable us to predict what is coming. We may not predict with certainty, but nevertheless a law is designed to aid in predicting. Prediction is not description. But when one is able to predict by use of laws based on experience (but not a *summary* of experience) he has one of the essential characters of explanation; the world of sense experience begins to take on the character of intelligibility. Isolated facts and a sheer account of them (if that is possible)

give us an unintelligible world. One may account *for* them by explanation, one factor of which is predictability of the future by way of laws—all of which transcend the given in that they have future referents which are neither experienced nor experienceable in a present sense datum. Intelligibility is entirely lacking in description *per se*. Although the predictive potency of statements is not description, it is essential to explanation.

III. PRIMITIVE EXPERIENCE AND EXPLANATION

As explained above, the other factor essential to the intelligibility of nature (or to explanation) is what I prefer to call a *native acquaintance* with the elements into which a situation can be analyzed. This analysis may be actual, but in most cases it is an ideal abstraction. This *native acquaintance* might well be called a *primitive subintellectual experience* (in the sense that more complex experiences are analyzable into it or deducible from it). This may seem like nonsense, but a little consideration will disabuse our minds of this conclusion.

Ernst Mach, who rebelled openly against traditional metaphysical theories of explanation, nevertheless has a theory of explanation which he supports adequately, not by sheer verbalization, but by historical facts, as we shall see later. He writes:

> "When once we reach the point where we are everywhere able to detect the *same* few simple elements, combining in the ordinary manner, then they appear to us as things that are familiar; we are no longer surprised, there is nothing new or strange to us in the phenomena, we feel at home with them, they no longer perplex us, they are *explained*" (5, p. 7).
>
> "The object that Archimedes and his successors sought to accomplish . . . , consists in the endeavor to reduce the more complicated case of the lever to the simpler and apparently self-evident case, to *discern* the simpler in the more complicated, or *vice versa*. In fact, we regard a phenomenon as explained, when we discover in it known simpler phenomena" (5, p. 19).

We will find that throughout the history of science men such as Archimedes, del Monte, Stevinus, Galileo, Daniel Bernoulli, Nemorarius, and others, who sought to establish mechanics on a sound basis, always tried to find that human experience which is somehow primitive and could, therefore, serve as a principle of explanation. (Of course, the science of mechanics, as we know, serves as the basis and as the prototype of many other branches of science.) This human experience which they sought *was not a purely intellectual experience*. It was, rather, kinesthetic experience, such as sensations from the muscles, tendons, and joints, coupled possibly with the sense of balance. They wanted that experience in which any person could both *feel* (or sense) and *see* (or discern) what is to follow. *E.g.*, some held that equal weights suspended from the lever arms at equal lengths from the (center) fulcrum would balance each other, and that nothing (by way of motion) would happen (they would be in a state of equilibrium). These contenders (5, Chaps. 1, 2) believed that this case is the very principle of statics which, in turn, is the basis for an *understanding* of dynamics. Others, of course, held that the balanced lever arms is a special case of dynamics.

These scientists were not asking the questions: Which situation will we accept arbitrarily as the conventional one? Which is simplest? Which is the one in terms of which we can *describe* other phenomena? They looked for kinesthetic sense experiences in terms of which, when made explicit to the mind, the world could become intelligible, or in terms of which their science would be significantly true of the world in the

sense that phenomena would then be *explained*. This will always be a part of the meaning of explanation. Explanatory experiences are not altogether intellectual. They are also primitive and necessary.

It is reasonable to believe that we have the experiential basis for an explanation of mechanics in some such experience as that of walking or lifting. In the former we continually rise and fall forward and, in a sense, maintain our balance. In the latter, we change the line of gravitational pull on our body and at the same time use the principle of the lever. Some such primitive experiences as these bridge the gap between mechanics as an abstract deductive science on the one hand and the practical application of that science on the other hand. These primitive, sub-intellectual experiences must be made explicit to the mind before an abstract, deductive, practical science can emerge. These experiences, when made explicit, or when we become intellectually conscious of them, constitute the basis of *explanation*.[1]

Surely the logical positivists (logical empiricists) are not objecting to this non-metaphysical type of explanation. For here we have shown what in principle is necessary if we are to have explanation and also a sound experiential basis for it. What positivists (and most scientists who have considered the matter) object to is a metaphysical basis for explanation. What does *metaphysical* mean to them? It means that an insurmountable logical breech has been set up between 'explanation' and things directly experiencable and waiting to be explained. Professor Herbert Feigl writes:

> "The locus of 'necessity' in the light of which the facts appear when explained lies . . . , exclusively in the relation of logical implication between premises and conclusions. This insight enables us to avoid the reductive fallacy according to which science is nothing more than description. For science pure and simply gives merely an *account of* but does not *account for* the observed phenomena in the way just elucidated" (3, 4).

Similarly Hans Reichenbach, in his well-taken new book, *The Philosophical Foundations of Quantum Mechanics* (8), is careful to show that scientific theory and interpretation must be confined to what he calls 'interphenomena' (*i.e.*, those events lying, spatially and temporally, between phenomena directly experienced but which, nevertheless, call for explanation). These interphenomena must be presented in such a way that they connect an earlier observed phenomenon, p_1, with a later observed phenomenon, p_2. These interphenomena are not presumably metaphysically different in structure or in reality than the phenomena (observed events). They connect phenomena logically and thereby justify deductive sciences.

I would go one step further, not in the direction of metaphysical abstractions which are emotionally satisfying but lead to logical gaps in science, but in the direction of more universal and more primitive, sub-logical experiences so as to justify our quest for explanation by finding in experience that which, when made explicit to us, bridges the logical gap between events whose relationship lies there to be explained. I would contend that Reichenbach's most worthy distinction between phenomena (observable events) and interphenomena must be pursued further in order to justify our quest for explanation without which laws and the predictive potency of scientific statements, and finally sci-

[1] Here we do not have space to show how these primitive, sub-intellectual experiences become explicit to the mind. The reader will find excellent suggestions in G. H. Mead's book (6), Supplementary Essays I and II.

ence itself, would be insignificant. But these interphenomena, if intelligible and if they bridge the logical gap between spatially and temporally separated experienced and experienceable phenomena, must be formulated in terms of these primitive experiences described above. In conclusion, the description and analysis of experienced phenomena will obviously never give us laws, theories, or logical connection between events. Consequently they will never give us explanation.

REFERENCES

1. CARNAP, R. Philosophy and logical syntax. In *Language, truth and logic* (A. J. Ayer, Ed.). New York: Oxford Univ. Press, 1936.
2. COMTE, A. The law of the three stages. In *Anthology of modern philosophy* (D. S. Robinson, Ed.). New York: Crowell, 1931. Pp. 697–711.
3. FEIGL, H. Logical empiricism. In *Twentieth century philosophy* (D. D. Runes, Ed.). New York: Philosophical Library, 1943, Part I, Chap. 1.
4. ——. Rejoinders and second thoughts. PSYCHOL. REV., 1945, 52, 278–281.
5. MACH, E. *The science of mathematics* (trans. by J. T. McCormack). La Salle: Open Court, 1942.
6. MEAD, G. H. *The philosophy of the present*. Chicago, London: Open Court, 1932.
7. PEARSON, K. *Grammar of science*. New York: Scribners, 1892.
8. REICHENBACH, H. *The philosophical foundations of quantum mechanics*. Berkeley: University of California Press, 1944.

OPERATIONISM AND PSYCHOLOGICAL THEORY: A NOTE

BY W. C. H. PRENTICE

The Johns Hopkins University

The recent symposium on operationism has had the effect of clarifying a large number of issues which have needed airing for several years. The authors have not, however, taken Question 7 seriously enough.[1] No one of them seems really worried that positivism may, to use a well-worn phrase, turn into negativism by stifling the development of new and useful theories. If operationists deny that they have ever tried to 'legislate' on the procedures of others, that is well and good; but their discussions and criticisms have produced a social climate in which the psychological theorist may hesitate to present theories which contain non-operational definitions.

A recent case in point is the revival by Spence[2] of an old controversy concerning 'continuous' *vs.* 'non-continuous' learning. Stated in terms which are not defined operationally but which are nevertheless intelligible to literate people, the basic question of this controversy runs as follows: "Does the mere presence of stimuli relevant to the solution of a problem affect the course of learning whether or not the subject becomes aware of their relevance?"

The 'continuity theorists'[3] say yes. According to their view, the mere physiological reception of a stimulus cue is sufficient to insure that associations will be set up and response tendencies determined. If now the relevance of such a cue be reversed (*e.g.*, the bright alley become 'correct' and be rewarded whereas it was previously 'incorrect' and led to punishment) the difficulty of learning the new problem will be increased.

The 'non-continuity theorists'[4] say no. If they are correct, no change in the relevance of cues to the true (or Experimenter's) solution can affect the course of S's learning unless the reversal is experienced as violating a 'hypothesis' already set up by S.

There are difficulties with both theories. But the important fact for the present discussion is that 'operationism' in its most radical form has served to becloud the issues.

The key to the whole controversy is the fact that Krechevsky,[5] impressed by

[1] Symposium on operationism. PSYCHOL. REV., 1945, 52, 241–294. Question 7 reads as follows: 7. Must operationists in psychology relegate theorizing of all sorts to the limbo of metaphysics? Bridgman in physics is perfectly aware of the value of theories as long as they are in keeping with his operational requirements. The Gestaltists, particularly Köhler and Koffka, have repeatedly attacked positivism (an identical twin of operationism), reproaching it for its (alleged) opposition to theoretical construction. C. C. Pratt (*Logic of Modern Psychology*, pp. 147–154) on the basis of his operationism maintains that all theoretical explanation is circular or tautological. Köhler (*Dynamics in Psychology*, pp. 107–125) holds a strictly opposite view. Which position is the most adequate for psychological research?

[2] Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. exp. Psychol.*, 1945, 35, 253–266.

[3] Principally Spence, T. L. McCulloch, and J. C. Pratt (see references in Spence, *op. cit.*) but supported by Hull.

[4] Principally I. Krechevsky and M. Haire (see references in Spence, *op. cit.*) but supported in part by Lashley and Köhler in various writings.

[5] Krechevsky, I. 'Hypotheses' vs. 'chance' in the pre-solution period in sensory discrimination learning. *Univ. Cal. Publ. Psychol.*, 1932, 6, 27–44.

the systematic nature of animal behavior during the 'pre-solution period' in learning experiments, proposed that the animals were not responding mechanically and in random fashion but were behaving in accordance with 'hypotheses.' But, and here is the crux of the matter, Krechevsky felt constrained to define the term 'hypothesis' operationally, and he defined it in terms of the animals' systematic responses themselves. The result is that the term 'hypothesis' immediately ceased to have a valid *raison d'être.*

The mere fact that Krechevsky chose the word 'hypothesis' suggests that the animals' systematic responses led him toward the belief that animal learning is at least partially governed by symbolic processes analogous to those in man. For in man a *hypothesis* is a symbolic solution which *precedes* attempts at verification on the manipulative or practical level. The notion that something similar may play a part in the activities of rats is a perfectly good theory and is supported by the systematic behavior of the animals during the 'pre-solution' period. But Krechevsky defeated his own cause by his overzealous operationism, and Spence has added to the confusion by insisting that 'hypothesis' and 'pre-solution period' are meaningless unless they are defined in terms of the animals' responses. Which is clearly not true.

At the human level we know that problems can be solved without overt behavior. When overt behavior is necessary to satisfy an Experimenter, how is 'solution' to be defined? Did it occur when the subject had the proper insight which enabled him to solve the problem? Or did it only occur when he proved to the Experimenter that he had had the proper insight? Suppose the 'continuity' *vs.* 'non-continuity' experiments to be carried out with human subjects: reversal after the insight (which was private) but before demonstration of the insight to E will certainly increase learning time. But the test of 'non-continuity' must be reversal *before insight.* The human subject, moreover, may have multiple hypotheses. While he is trying out one hypothesis (*e.g.*, that the 'correct' door follows a double alternation pattern) he may also be noticing whether the correct one is always the lighted one (without, however, making overt systematic responses to the 'lighted' or 'non-lighted' characteristic). Who shall say that these things do not occur in rats? Any psychologist will admit that such mechanisms cannot be assumed until objective criteria are available, and the search for such criteria must continue, but the attempt by Spence *et al.* to establish an alternative theory by insisting on definitions which do not define only confuses the issue.

Two difficulties may be noted. First, the very term 'hypothesis' is misleading when it has been defined to mean a kind of behavior. If hypotheses in rats were what they are in man, the hypothesis would necessarily *precede* the behavior which indicated its presence, and lack of that behavior could not disprove the existence of the hypothesis. Second, and similarly, the operational definition of 'pre-solution period' has made the 'non-continuity' theory apparently subject to tests which actually are not crucial. In terms of the accepted definitions, Spence and his followers surely have the better of the argument; but the fact remains that those accepted definitions have prevented the 'non-continuity theory' as stated by Lashley from ever receiving an adequate test.

Operationists will reply that the Krechevsky-Spence controversy springs from *bad* operational definitions, not from operationism as such. Of course. But so long as psychologists feel forced to supply operations where none exist, they cannot escape the dilemma: give

up your concept (at least, do not publish it) or supply inadequate definitions which confuse scientific thinking despite a spurious appearance of precision.

We are still without operations which can define 'noticing the stimulus cues' or 'having a (symbolic) hypothesis' in animals, but the concepts are surely not meaningless. The implied insistence that "what is not operationally definable has no place in science" has forced the dilution of a good theory to a strength potable to the operationists; but it has thereby destroyed precisely the flavor which was valuable.

Here is the common ground of positivism and negativism. Here is where they meet. In the normal course of events, psychologists would have worked out the tests, observations and experiments necessary to objectify such concepts as 'hypothesis' and 'solution' since that is the normal procedure of science. If no technique appeared in the long run, the concept would have been dropped as unfruitful (but not *meaningless*). Actually, however, the operationist kibitzer has seemed to say, "You cannot even talk about such things in science unless you define them in operational terms. And I mean now." So the possibility of solving some basic problems of learning has been snowed under by meaningless experiments and more meaningless argument. The positivistic pressure has completely negated the force of Lashley's and Krechevsky's theories.